**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/56, 1/21, 9/28
A23K 3/03

A1

(11) International Publication Number: WO 92/19744

(43) International Publication Date: 12 November 1992 (12.11.92)

(21) International Application Number: PCT/FI92/00138

(22) International Filing Date: 4 May 1992 (04.05.92)

(30) Priority data:
695,214 6 May 1991 (06.05.91) US

(71) Applicant: VALIO, FINNISH CO-OPERATIVE DAIRIES' ASSOCIATION [FI/FI]; PL 390, SF-00101 Helsinki (FI).

(72) Inventors: KOIVULA, Teija ; Porvoonkatu 16 A 36, SF-00510 Helsinki (FI). SIBAKOV, Mervi ; Ylipalontie 21 C, SF-00670 Helsinki ( ). PALVA, Ilkka ; Välitalontie 87 A, SF-00660 Helsinki (FI).

(74) Agent: OY JALO ANT-WUORINEN AB; Laivurinrinne 2 A, SF-00120 Helsinki (FI).

(81) Designated States: AT (European patent), AU, BE (European patent), BR, CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GB (European patent), GR (European patent), HU, IT (European patent), JP, LU (European patent), MC (European patent), NL (European patent), NO, SE (European patent).

Published
*With international search report.*

(54) Title: ACID α-AMYLASE

(57) Abstract

The invention is related to the recombinant production of acid α-amylase, DNA sequences encoding such acid α-amylase, vectors containing such DNA, and the expression of such DNA in a recombinant host cell. The invention is further related to a method of feed and fodder preservation using the hosts of the invention.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| AT | Austria | FI | Finland | ML | Mali |
| AU | Australia | FR | France | MN | Mongolia |
| BB | Barbados | GA | Gabon | MR | Mauritania |
| BE | Belgium | GB | United Kingdom | MW | Malawi |
| BF | Burkina Faso | GN | Guinea | NL | Netherlands |
| BG | Bulgaria | GR | Greece | NO | Norway |
| BJ | Benin | HU | Hungary | PL | Poland |
| BR | Brazil | IE | Ireland | RO | Romania |
| CA | Canada | IT | Italy | RU | Russian Federation |
| CF | Central African Republic | JP | Japan | SD | Sudan |
| CG | Congo | KP | Democratic People's Republic of Korea | SE | Sweden |
| CH | Switzerland | | | SN | Senegal |
| CI | Côte d'Ivoire | KR | Republic of Korea | SU | Soviet Union |
| CM | Cameroon | LI | Liechtenstein | TD | Chad |
| CS | Czechoslovakia | LK | Sri Lanka | TG | Togo |
| DE | Germany | LU | Luxembourg | US | United States of America |
| DK | Denmark | MC | Monaco | | |
| ES | Spain | MG | Madagascar | | |

# Acid α-Amylase

## Field of the Invention

The present invention is directed to purified genetic and protein sequences encoding the acid α-amylase protein of *B. acidocaldarius*, recombinant constructs and hosts transformed therewith that are capable of expressing such sequences, and the use of such genetic sequences for the expression of acidic αamylase, and the use of the recombinantly-produced acid α-amylase protein for industrial and agricultural applications such as, for example, starch liquefaction, and silage preparation.

## Brief Description of the Background Art

*Lactobacillus* and *Bacillus* strains offer many potential advantages in the production of cloned gene products, as compared with *Escherichia coli*. First, they are non-pathogenic and do not synthesize endotoxins. Second, many of the gene products are secreted into the growth medium, in contrast to *E. coli*, which retains most of the proteins due to the presence of an outer membrane. Third, they have been widely used for production of enzymes in large-scale fermentation processes. While *Lactobacillus* is not an industrial enzyme producer, it is highly used in the food and animal feed processing industry.

One important highly desirable exoenzyme secreted in large amounts by *Bacilli* is α-amylase. α-Amylase is used in a variety of industrial applications, for

example, starch liquefaction. However, the industrial use of this enzyme is limited in many cases by its sensitivity to acidic conditions. That is, most forms of the enzyme are not enzymatically stable at acidic pH's.

At least four acidophilic and acid-stable α-amylase isozymes have been reported in different strains of *B. acidocaldarius* (Kanno, M., *Agric. Biol. Chem. 50*:23-31 (1986); Buonocore, V. *et al., J. Bacteriol. 128*:515 (1976); Boyer, E.W.. *et al., Starch/Stärke 31*:166 (1979); and Uchino, F., *Agric. Biol. Chem. 46*:7 (1982). An "acidic" α-amylase is characterized as being relatively stable to heat under acidic conditions. For example, the acid α-amylase's produced by *Bacillus acidocaldarius* strain A2 retains more than 79-90% of its activity after 30 min of incubation at pH 2.0 (70°C) and at pH 4.5 (90°C) in the absence of substrate (Kanno, M., *Agric. Biol. Chem. 50*:23-31 (1986).

α-Amylase has also been isolated from *B. amyloliquefaciens* (Ingle *et al., Adv. Appl. Microbiol. 24*:257-278 (1987)). This enzyme has an $M_r$-value of about 50,000 daltons and has been sequenced *(Takkinen et al., J. Biol. Chem 258*: 1007-1013 (1983); Chung *et al., Biochem. J. 185*:387-395 (1980)). The expression of this enzyme in *B. subtilis* has been reported ("Expression and Regulation of the *Bacillus amyloliquefaciens* α-amylase gene in *B. subtilis*," P.Kallio, Ph.D. dissertation, University of Helsinki, 1987).

Acidic amylase sequences are especially desirable for industrial purposes, where pH values may fall (meaning the medium becomes more acidic) during a

desired reaction due to the natural catalytic action of the enzymes therein. For example, a process strain that possesses an acid α-amylase activity would be valuable for the *Lactobacillus* strains as lactic acid bacteria are widely used for preservation of many starch containing raw materials (e.g., cereal grains, edible roots and crop residues).

Although such approaches and applications have been highly desired in the art (*Applied Environment. Microbiol.* 55:2130-2137 (1989)), they have not been possible to obtain. The main obstacle was the lack of an α-amylase enzyme that is not inactivated by the low growth pH of *Lactobacillus*. So far, no genes of bacterial origin, coding for an acid α-amylase enzyme, have been characterized and the isolation of large quantities of such acid α-amylases from natural sources is relatively expensive. This has precluded the wide-spread adoption of such enzymes in industry. Accordingly, a need exists for a cost-efficient recombinant source of a highly stable acid α-amylase.

## SUMMARY OF THE INVENTION

Recognizing the importance to the feed industry of a host that would be capable of providing an acid-stable α-amylase activity during the food preservation period, and cognizant of the lack of hosts capable of providing such enzymatic activity in the currently utilized processes, the inventors have investigated the properties of the acid α-amylase from *B. acidocaldarius*. The inventors desired to generate a new *Lactobacillus* host, such host having a novel

property that would be advantageous in the food and feed industry.

These efforts have culminated in the development of novel recombinant forms of acid α-amylase, hosts expressing such acid α-amylase, and highly improved processes for the preservation of food utilizing such hosts.

According to the invention, there are first provided a polynucleotide, such polynucleotide providing the coding sequence of an acid α-amylase.

According to the invention, there are also provided recombinant vectors, such vectors providing a host expressible form of acid α-amylase.

According to the invention, there are further provided host cells transformed with polynucleotides and/or vectors that are capable of expressing recombinant forms of acid α-amylase.

According to the invention, there are further provided methods for producing the genetically engineered or recombinant protein acid α-amylase of the invention using such hosts.

According to the invention, there are further provided methods for the utilization of the acid α-amylase enzyme of the invention.

## DESCRIPTION OF THE FIGURES

Figure 1. SDS-PAGE of the $(NH_4)_2SO_4$ precipitated proteins. Lanes 1 and 2 represent Coomassie Brilliant Blue R stained gel and zymogram, respectively. The arrows show the positions of the bands corresponding amylase activity. Molecular weight markers (Pharmacia) are indicated on the left.

Figure 2. Effect of pH on the activity of *B. acidocaldarius* amylase(s). Enzyme activity in 100 µl of the 0.5 M sodium acetate, pH 5.0 eluted proteins of *B. acidocaldarius* cultures was measured using Phadebas® amylase test at 60°C in 0.1 M citric acid, 0.2 M $Na_2HPO_4$.

Figure 3. The nucleotide sequence of the *B. acidocaldarius* acid α-amylase gene. The putative -35 and -10 regions (bases 126 and 148, respectively) of the amylase promoter are underlined. The potential signal sequence cleavage site (base 273) is indicated by an arrow. The N-terminal sequence (base 1594) of the 90 kd protein is underlined with a dotted line. The two alternative N-termini of the signal peptide are indicated by an asterisk.

## DESCRIPTION OF THE PREFERRED EMBODIMENTS

### I. Definitions

In the description that follows, a number of terms used in recombinant DNA (rDNA) technology are extensively utilized. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

Process strain. A "process strain" is a bacterial strain that causes or contributes to a desired change in a biotechnical process.

Biotechnical process. A biotechnical process is a process that depends upon the production of bacteria or biomass for the achievement of a desired

chemical transformation of the process medium, such bacteria or biomass providing a component or environment necessary for such conversion.

Fodder. "Fodder" is anything fed to domesticated animals, and especially, coarse food for cattle, horses or sheep.

Forage. "Forage" is food for domesticated animals that is taken by browsing or grazing.

Silage. "Silage" is fodder converted into succulent feed for livestock through processes of anaerobic acid fermentation (as occurs in a silo).

Gene. A DNA sequence containing a template for a RNA polymerase. The RNA transcribed from a gene may or may not code for a protein. RNA that codes for a protein is termed messenger RNA (mRNA).

A "complementary DNA" or "cDNA" gene includes recombinant genes synthesized by reverse transcription of mRNA and from which intervening sequences (introns) have been removed.

Genetic sequence. As used herein, the term "genetic sequences" is intended to refer to a nucleic acid molecule (preferably DNA). Accordingly, "acid α-amylase genetic sequences" are DNA or RNA sequences that possess a nucleotide sequence that encodes the acid α-amylase protein.

Promoter. The term "promoter" as used herein refers to a module or group of modules which, at a minimum, provides a binding site or initiation site for RNA polymerase action that is sufficient to initiate transcription of the operably linked in a desired host. A promoter is generally composed of multiple operably linked genetic elements termed herein "modules."

Promoter Module. The term "module" as in "promoter module" refers to a genetic transcriptional regulatory element which provides some measure of control over the transcription of operably linked coding sequences or other operably linked modules.

Each module in a promoter can convey a specific piece of regulatory information to the host cell's transcriptional machinery. At least one module in a promoter functions to position the start site for RNA synthesis. Other promoter modules regulate the frequency of transcriptional initiation. Typically, modules which regulate the frequency of transcriptional initiation are located upstream of (i.e., 5' to) the transcriptional start site, although such modules may also be found downstream of (i.e., 3' to) the start site.

The term "target module," as used herein, refers to a transcriptional regulatory element which confers the ability to respond to enhancer gene activity (i.e., such as the protein or peptide encoded by an enhancer gene) on a promoter which otherwise would not respond, or would respond less efficiently, to such enhancer gene activity.

The term "initiation module" refers to a promoter module which is required to initiate transcription of operably linked genes with RNA polymerase. In prokaryotic promoters, initiation modules are usually located at about -10 and -35 nucleotides from the start site of transcription.

By "hybrid promoter" is meant a promoter in which an initiation module is operably linked to a heterologous target module. A target module which is heterologous to an initiation module is a target

module which is not found naturally operably linked to this initiation module in the host cell.

Operable linkage. An "operable linkage" is a linkage in which a sequence is connected to another sequence (or sequences) in such a way as to be capable of altering the functioning of the sequence (or sequences). For example, a protein encoding sequence which is operably linked to the hybrid promoter of the invention places expression of the protein encoding sequence under the influence or control of the regulatory sequence. Two DNA sequences (such as a protein encoding sequence and a promoter region sequence linked to the 5' end of the encoding sequence) are said to be operably linked if induction of promoter function results in the transcription of the protein encoding sequence mRNA and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the expression regulatory sequences to direct the expression of the mRNA or protein. Thus, a promoter region would be operably linked to a DNA sequence if the promoter were capable of effecting transcription of that DNA sequence.

Cloning vector. A "cloning vector" is a plasmid or phage DNA or other DNA sequence which is able to replicate autonomously in a host cell, and which is characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without loss of an essential biological function of the vector, and into which DNA may be spliced in order to bring about its replication and cloning. The cloning vector may

further contain a marker suitable for use in the identification of cells transformed with the cloning vector. Markers, for example, are erythromycin and kanamycin resistance. The term "vehicle" is sometimes used for "vector."

Expression vector. An "expression vector" is a vector similar to a cloning vector but is capable of expressing a structural gene which has been cloned into the expression vector; after transformation of the expression vector into a host. In an expression vector, the cloned structural gene (any coding sequence of interest) is placed under the control of (i.e., operably linked to) certain control sequences which allow such gene to be expressed in a specific host. In the expression vector of the invention, a desired structural gene is operably linked to the hybrid promoter of the invention. Expression control sequences will vary, and may additionally contain transcriptional elements such as termination sequences and/or translational elements such as initiation and termination sites.

The expression vectors of the invention may further provide, in an expression cassette other than the one providing the hybrid promoters of the invention, sequences encoding a desired enhancer gene. In a preferred embodiment, such enhancer gene would be the enhancer gene which encodes the protein which regulates the target module of the hybrid promoter.

Functional Derivative. A "functional derivative" of a molecule, such as a nucleic acid or protein, is a molecule which has been derived from a native molecule, and which possesses a biological activity (either functional or structural) that is substan-

tially similar to a biological activity of the native molecule, but not identical to the native molecule. By a functional derivative of a nucleic acid sequence (herein the "second sequence") that encodes a protein (herein a "first" sequence) is meant (1) a nucleic acid sequence that does not possess the exact same nucleotide sequence as the first sequence but which encodes the same amino acid sequence; and, a nucleic acid sequence that does not encode the same amino acid sequence but which encodes a protein with biochemical properties (such as, for example, enzymatic stability in acidic conditions) equivalent to (i.e., the same as or slighly different from) the protein encoded by the first nucleic acid sequence.

A functional derivative of a protein is a protein that retains a desired function of the "parent" protein from which it is derived. A functional derivative of a protein may or may not contain post-translational modifications, such as covalently linked carbohydrate, depending on the necessity of such modifications for the performance of a specific function. The term "functional derivative" is intended to include the "fragments," "variants," or "chemical derivatives" of a molecule.

As used herein, a molecule is said to be a "chemical derivative" of another molecule when it contains additional chemical moieties not normally a part of the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of mediating such effects are

disclosed in *Remington's Pharmaceutical Sciences* (1980). Procedures for coupling such moieties to a molecule are well known in the art.

Fragment. A "fragment" of a molecule such as a nucleic acid or protein is meant to refer to a molecule which contains a portion of the complete sequence of the native molecule.

Variant. A "variant" of a molecule such as a nucleic acid or protein is meant to refer to a molecule substantially similar in structure and biological activity to either the entire molecule, or to a fragment thereof, but not identical to such molecule or fragment thereof. A variant is not necessarily derived from the native molecule itself. A gene allele is an example of what is meant by a variant nucleic acid sequence and an enzymatic isozyme is an example of what is meant by a variant enzyme sequence. Thus, provided that two molecules possess a similar biological activity or function that characterizes molecules as a 'family', they are considered variants as that term is used herein even if the composition or secondary, tertiary, or quaternary structure of one of the molecules is not identical to that found in the other, or if the sequence of nucleic acid (or amino acid residues) is not identical, or if the synthesis of one of the variants did not derive from the other. Thus the invention intends to include all variants (genomic alleles and enzymatic isozymes) of the acid α-amylase of the invention.

II. Genetic Engineering of Acid α-amylase

Prior to the invention, the cloning of the acid amylase gene was not thought to be possible due to the lack of cross-reacting antibodies, the lack of sufficient protein to obtain such antibodies, the lack of a reliable probe and the lack of protein sequence data from which to derive such probe.

The available antisera or gene probes of other known bacterial amylases (for example, those of *B. amyloliquefaciens* strain A2 as described by Kanno, M. *Agric. Biol. Chem. 50*:23-31 (1986), *B. licheniformis* as described by Siato, N., *Arch. Biochem. Biphys. 155*:290 (1973) and *B. subtilis*) does not provide any crossreactivity or positive hybridization when tested with material derived from *B. acidocaldarius*. When the amylase preparation from *B. amyloliquefaciens* was analyzed by zymography, it revealed an amylase band of molecular weight 150,000-200,000 daltons. This is much higher than previously reported for other bacterial amylases. Due to the very low initial yield and large size of the acid amylase, it was not possible to determine the N-terminal sequence of this protein by standard protein chemistry methods in order to obtain an oligonucleotide probe.

Thus it was necessary to identify clones to the *B. acidalcarius* acid amylase without having available (1) antibodies that cross-reacted with the protein or (2) Protein or DNA sequence information upon which to design a oilgonucleotide probe.

To solve this problem, a new method for the identification of acid amylase protein and clones was developed. A 90,000 dalton peptide was found in

protein extracts of *B. acidocaldarius*, that had a very low specific amylase activity. An assumption was made that this peptide would be sufficiently related to the large acid amylase protein to allow selection of clones to the acid amylase protein, using probes designed from the sequence of the smaller peptide. This was not a risk-free assumption as the low level "amylase" activity displayed by the 90,000 dalton peptide could easily have been a side activity of a different type of enzyme, and not related to the acid amylase enzyme of the invention *per se*. Thus, there was no way of confirming that the assumption upon which the cloning of the enzyme of the invention was based was correct until the entire cloning and sequencing was complete.

The process for genetically engineering the acid α-amylase sequences of the invention is facilitated through the cloning of genetic sequences which are capable of providing specific protein encoding sequences. Genetic sequences which are capable of providing protein encoding sequences may be derived from genomic DNA, synthetic DNA, cloned DNA and combinations thereof. The preferred species source of the acid α-amylase of the invention is *B. acidocaldarius*, although any source of an acid α-amylase may be used.

Genetic (protein encoding) genomic DNA will not contain introns in prokaryotes, although it may contain spacers between transcriptional units. As outlined below, such genomic DNA may be obtained in association with the 5' promoter region and/or the 3' transcriptional termination region if desired. Further, such genomic DNA may be obtained in

association with the genetic sequences which encode a 5' non-translated region of the desired mRNA and/or with the genetic sequences which encode the 3' non-translated region. To the extent that a host cell can recognize the transcriptional and/or translational regulatory signals associated with the expression of the mRNA and protein, and to the extent that such signals do not impede the expression of the recombinant acidic αamylase of the invention, then the 5' and/or 3' non-transcribed regions of the native gene, and/or, the 5' and/or 3' non-translated regions of the mRNA, may be retained and employed for transcriptional and translational regulation.

The coding sequences for the acid α-amylase of the invention is not rearranged by the native bacterial host prior to expression in such host. To obtain coding sequences for proteins whose genes are not rearranged prior to expression, genomic DNA can be extracted and purified from any cell of any host which carries the coding sequence, whether or not the cell expresses the protein. Such extraction of genomic DNA can be performed by means well known in the art (for example, see *Guide to Molecular Cloning Techniques*, S.L. Berger *et al.*, eds., Academic Press (1987)).

Alternatively, nucleic acid sequences that encode a desired protein can also be obtained by DNA transcribed from mRNA specific for that protein. mRNA can be isolated from any cell which produces or expresses the protein of interest and used to produce cDNA by means well known in the art (for example, see *Guide to Molecular Cloning Techniques*, S.L. Berger *et al.*, eds., Academic Press (1987)). Preferably, the mRNA preparation used will be enriched in mRNA coding

for the desired protein, either naturally, by isolation from a cells which are producing large amounts of the protein, or *in vitro*, by techniques commonly used to enrich mRNA preparations for specific sequences, such as for example, gel electrophoresis, sucrose gradient centrifugation.

To prepare DNA for cloning into a cloning vector or an expression vector, a suitable DNA preparation (either genomic DNA or cDNA) is randomly sheared or enzymatically cleaved, respectively. Such DNA can then be ligated into appropriate vectors to form a recombinant gene (either genomic or cDNA) library.

A DNA sequence encoding a protein of interest or its functional derivatives may be inserted into a cloning vector or an expression vector in accordance with conventional techniques, including blunt-ending or staggered-ending termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and ligation with appropriate ligases. Techniques for such manipulations are disclosed by Maniatis, T., *et al., Molecular Cloning (A Laboratory Manual)*, Cold Spring Harbor Laboratory, second edition, 1989), and are well known in the art.

Libraries containing clones encoding a desired protein or a desired transcriptional regulatory element may be screened and a desired clone identified by any means which specifically selects for the DNA of interest. For example, if a clone to an acid α-amylase is desired, such a clone may be identified by any means used to identify acid α-amylase protein or mRNA, including, for example, a) by hybridization with

an appropriate nucleic acid probe(s) containing a sequence(s) specific for the DNA of this protein, or b) by hybridization-selected translational analysis in which native mRNA that hybridizes to the clone in question is translated *in vitro* and the translation products are further characterized, or, c) if the cloned genetic sequences are themselves capable of expressing mRNA, by immunoprecipitation of a translated protein product produced by the host containing the clone.

Oligonucleotide probes specific for a desired protein can be used to identify a desired clone. Such probes can be designed from knowledge of the amino acid sequence of the desired protein. The sequence of amino acid residues in a peptide is designated through the use of the commonly employed three-letter or single-letter designations. A listing of these three-letter and one-letter designations may be found in textbooks such as *Biochemistry*, Lehninger, A., Worth Publishers, New York, NY (1970). As used herein, when the amino acid sequence is listed horizontally, the amino terminus is intended to be on the left end and the carboxy terminus is intended to be at the right end.

Because the genetic code is degenerate, more than one codon may be used to encode a particular amino acid. The peptide fragments are analyzed to identify sequences of amino acids that may be encoded by oligonucleotides having the lowest degree of degeneracy. This is preferably accomplished by identifying sequences that contain amino acids that are encoded by only a single codon.

Although occasionally an amino acid sequence may be encoded by only a single oligonucleotide sequence, frequently the amino acid sequence may be encoded by any of a set of similar oligonucleotides. Importantly, whereas all of the members of this set contain oligonucleotide sequences that are capable of encoding the same peptide fragment and, thus, potentially contain the same oligonucleotide sequence as the gene which encodes the peptide fragment, only one member of the set contains the nucleotide sequence that is identical to the exon coding sequence of the gene. Because this member is present within the set, and is capable of hybridizing to DNA even in the presence of the other members of the set, it is possible to employ the unfractionated set of oligonucleotides in the same manner in which one would employ a single oligonucleotide to clone the gene that encodes the peptide.

Therefore, using the genetic code, one or more different oligonucleotides can be identified from the amino acid sequence, each of which would be capable of encoding the desired protein. The probability that a particular oligonucleotide will, in fact, constitute the actual protein's encoding sequence can be estimated by considering abnormal base pairing relationships and the frequency with which a particular codon is actually used (to encode a particular amino acid) in the host cell. Such "codon usage rules" are disclosed, for example, by Lathe, R., *et al., J. Molec. Biol. 183*:1-12 (1985). Using the "codon usage rules" of Lathe, a single oligonucleotide sequence, or a set of oligonucleotide sequences, that contain a theoretical "most probable" nucleotide

sequence capable of encoding the acid α-amylase of the invention is identified.

In addition to the codon usage rules, oligonucleotide design can may utilize the use of deoxyinosine at ambiguous codon positions. This approach is particularly useful when the required DNAs sequence is derived from a poorly characterized organism like *B. acidocaldarius* (Takahaski *et al., Proc. Natl. Acad. Sci. USA 82*:1931-1935 (1985)).

The suitable oligonucleotide, or set of oligonucleotides, that are capable of encoding a fragment of the desired gene (or which is complementary to such an oligonucleotide, or set of oligonucleotides) may be synthesized by means well known in the art (see, for example, *Synthesis and Application of DNA and RNA*, S.A. Narang, ed., 1987, Academic Press, San Diego, CA) and employed as a probe to identify and isolate the cloned gene by techniques known in the art. Techniques of nucleic acid hybridization and clone identification are disclosed by Maniatis, T., *et al. (In: Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY (1982)), and by Hames, B.D., *et al.*, in: *Nucleic Acid Hybridization, A Practical Approach*, IRL Press, Washington, DC (1985)), which references are herein incorporated by reference. Those members of the above-described gene library which are found to be capable of such hybridization are then analyzed to determine the extent and nature of the sequences which they contain.

To facilitate the detection of the desired encoding sequence, the above-described DNA probe may be labeled with a detectable group. Such detectable

group can be any material having a detectable physical or chemical property. Such materials have been well-developed in the field of nucleic acid hybridization and in general most any label useful in such methods can be applied to the present invention. Particularly useful are radioactive labels, such as $^{32}P$, $^{3}H$, $^{14}C$, $^{35}S$, $^{125}I$, or the like. Any radioactive label may be employed which provides for an adequate signal and has a sufficient half-life. The oligonucleotide may be radioactively labeled by means well-known in the art, for example, "nick-translation" and T4 DNA polymerase replacement synthesis.

Alternatively, polynucleotides are also useful as nucleic acid hybridization probes when labeled with a non-radioactive marker such as biotin, an enzyme or a fluorescent group.

Thus, in summary, the actual identification of acid α-amylase peptide sequences permits the identification of a theoretical "most probable" DNA sequence, or a set of such sequences, capable of encoding such acid α-amylase. By constructing an oligonucleotide complementary to this theoretical sequence (or by constructing a set of oligonucleotides complementary to the set of "most probable" oligonucleotides), one obtains a DNA molecule (or set of DNA molecules), capable of functioning as a probe(s) for the identification and isolation of clones directed to acid α-amylase.

The above discussed methods are, therefore, capable of identifying genetic sequences that are capable of encoding acid α-amylase, or fragments, variants and functional derivatives of acid α-amylase. In order to further characterize the cloned acid α-

amylase (or its fragment, variant or functional derivative) and especially, in order to produce recombinant acid α-amylase, it is desirable to express the proteins which the sequences encodes.

III. Expression of Proteins Using the Expression Vectors of the Invention

Expression of acid α-amylase protein allows the identification of clones that are capable of expressing the acid α-amylase protein of the invention, or fragments, variants or functional derivatives thereof. Characteristics unique to acid α-amylase that may be used to identify the acid α-amylase protein, fragment, variant or functional derivative include the ability to specifically bind acid α-amylase antibodies, the ability to elicit the production of acid α-amylase antibodies that are capable of binding to the native protein, and the ability to provide an enzymatic function specific to acid α-amylase, such as, for example, enzymatic stability at an acidic pH at which other "non-acidic" α-amylase enzymes are relatively unstable, among others.

In a preferred embodiment, amylase-negative bacterial strains are used as hosts. Especially preferred are amylase negative mutants of *Bacillus* or *L. plantarum* (which does not produce amylase by definition).

To express the acid α-amylase of the invention, transcriptional and translational signals recognizable by an appropriate host are necessary. C l o n e d

sequences encoding acid α-amylase, its fragments, variants or functional derivatives, obtained through the methods described above, and preferably in a double-stranded form, may be operably linked to sequences controlling transcriptional expression in an expression vector. Such sequences may be introduced into a host cell to produce recombinant acid α-amylase, its fragments, variants and functional derivatives.

In addition to transcriptional signals, it may also be advantageous to use translational and/or secretion signals derived from other bacterial sources provided that such translational and secretion signals are functional in the chosen host cell. This is discussed further *infra*.

According to the invention, any prokaryote host may be utilized. In a preferred embodiment, a member of the *Bacillus* or *Lactobacillus* genera are used as the host cell for expressing the desired protein of the invention. Such members include *B. subtilis, B. licheniformis, B. amyloliquefaciens, B. polymyxa, B. stearothermophilus, B. theroproteolyticus, B. coagulans, B. thuringiensis, B. megaterium, B. cereus, B. natto,* and, *B. acidocaldarius*. In an especially highly preferred embodiment, the host cell is *Lactobacillus*.

*Lactobacillus* species that are associated with silage and are expecially preferred as hosts in the processes of the invention are *L, plantarum, L. brevis, L. buchnerie, L. coryniformis, L. curvatus, L. casei, L. fermentum, L. acidophilus,* and *L. salivarius*.

A nucleic acid molecule, such as DNA, is said to be "capable of expressing" a a polypeptide like the acid α-amylase of the invention if the nucleic acid molecule contains expression control sequences that provide transcriptional regulatory information and such sequences are operably linked to the nucleotide sequence which encodes the acid α-amylase.

The precise nature of the regulatory regions needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribing and 5' non-translating (non-coding) sequences involved with initiation of transcription and translation respectively. Especially, such 5' non-transcribing control sequences will include a region which contains a promoter capable of functioning in the host cell.

Expression of a recombinant protein in prokaryotic hosts requires the use of regulatory regions functional in such hosts, and preferably prokaryotic regulatory systems. A wide variety of transcriptional and translational regulatory sequences can be employed, depending upon the nature of the prokaryotic host. Preferably, these regulatory signals are associated with a particular gene which is capable of a high level of expression in the host cell.

If desired, a fusion product of the desired protein may be constructed. For example, if the genetic sequence encoding a desired acid α-amylase (or its fragment, variant or functional derivative) does not possess a sequence encoding a signal sequence functional in a certain host, such signal sequence may be operably linked to the desired genetic sequence, thus allowing secretion of the protein from, or the

membrane compartmentalization of the protein in, the host cell. Such signal sequences may be designed with or without specific protease sites such that the signal peptide sequence is amenable to subsequent removal. Alternatively, the native signal sequence of a protein of the invention may be used.

In addition to using either a native signal sequence or a signal derived from some other bacteria, it is often useful to construct a signal sequence fusion sequence, utilizing only part of the native signal sequence. Potential fusion sites between two signal sequences are the hydrophobic regions or the helix breaker residue between the hydrophobic region and the C-terminal part of the signal seuqence. Examples of useful fusion signal sequences are described in Applicants' copending application, U.S. Application No. 07/377,450, filed July 10, 1989 and incorporated herein fully by reference.

Transcriptional initiation regulatory signals that can be operably linked to the proteins of the invention can be selected which allow for repression or activation, so that expression of the operably linked genes can be modulated in a specific manner.

Where the native expression control sequences signals do not function satisfactorily in a desired host cell, then sequences functional in the host cell may be substituted.

Once the vector or DNA sequence containing the construct(s) is prepared for expression, the DNA construct(s) is introduced into an appropriate host cell by any of a variety of suitable means known in the art. In a preferred embodiment, *Lactobacillus* are

transformed essentially as described in von Wright *et al., Appl. Enciron. Microbiol. 56*:2029-2035 (1990).

To transform a host cell with the DNA constructs of the invention many vector systems are available, depending upon whether it is desired to insert the genetic DNA construct into the host cell chromosomal DNA, or to allow it to exist in an extrachromosomal form. When it is desired to maintain a vector in an extrachromosomal form the vector should provide an origin of replication capable of functioning in the host.

Genetically stable transformants may also be constructed with vector systems, or transformation systems, whereby a desired protein's DNA is integrated into the host chromosome. Such integration may occur *de novo* within the cell or, in a most preferred embodiment, be assisted by transformation with a vector which functionally inserts itself into the host chromosome. For example, such vector may provide a DNA sequence element which promotes integration of DNA sequences in chromosomes. In a preferred embodiment, such DNA sequence element is a sequence homologous to a sequence present in the host chromosome such that the integration is targeted to the locus of the genomic sequence and targets integration at that locus in the host chromosome.

Cells that have stably integrated the introduced DNA into their chromosomes are selected by also introducing one or more markers which allow for selection of host cells that contain the expression vector in the chromosome, for example the marker may provide biocide resistance, e.g., resistance to antibiotics, or the like. The selectable marker gene

can either be directly linked to the DNA gene sequences to be expressed, or introduced into the same cell by co-transformation.

A sequence may also be incorporated into a plasmid or other vector capable of autonomous replication in the recipient host. Any of a wide variety of vectors may be employed for this purpose. Factors of importance in selecting a particular vector include: the ease with which recipient cells that contain the vector may be recognized and selected from those recipient cells which do not contain the vector; the number of copies of the vector which are desired in a particular host; and whether it is desirable to be able to "shuttle" the vector between host cells of different species.

In addition to antibiotic selection markers, one can (or should, if such antibiotic markers are undesirable in a specific process in which the hosts will be utilized) use natural selection markers based on, for example, nisin resistance, thymidine synthesis, lactose utilization or X-prolyl-dipeptidyl-aminopeptidase (*Appl. Environ. Microbiol.* 57:38-43 (1991).

After the introduction of the vector, recipient cells are grown in a selective medium, which selects for the growth of vector-containing cells. Expression of the cloned gene sequence(s) results in the production of the desired protein. This expression can take place in a continuous manner in the transformed cells, or in a controlled manner.

The expressed acid α-amylase, or fragment, variant or functional derivative thereof may be isolated and purified in accordance with the method

described herein and using techniques known in the art. However, in a highly preferred embodiment, no isolation or purification of the product is needed as the host is directly inoculated into the process medium.

The genetic sequences, protein sequences, vectors and methods of the invention are useful in many applications of the food and beverage industry. Especially, the hosts of the invention are useful when directly inoculated into foodstuff. The lactic acid bacteria of the invention may be utilized for the preparation of fermented food and beverages as lactic acid bacteria play an essential role in the preparation of such fermented food and beverages. *The Lactobacilli* of the invention may also be applied in silage preparation and as probiotics in human and animal health. The hosts containing the sequences and methods taught herein, together with, if necessary, current knowledge of the genetics and biochemistry of lactic acid bacteria, especially of *Lactococci* and *Lactobacillus*, may be used for the construction of bacterial argicultural process strains that have been optimized for different agricultural and food preparation applications using the sequences and methods of the invention. The genetics of Lactic acid bacteria are known in the art and are reviewed, for example, in *Biochimie 70*:No. 3 and No. 4 (1988) and in *FEMS Microbiol. Rev. 87*:No. 1 and No. 2 (1990).

For example, one optimization using the sequences and methods of the invention is the addition of heterologous acid $\alpha$-amylase activities, and especially hosts expressing such activities, to any process strain. Highly desirable hosts that may be utilized

as process strain host microorganisms include *B. subtilis, B. licheniformis, B. amyloliquefaciens, B. polymyxa, B. stearothermophilus, B. thermoproteolyticus, B. coagulans, B. thuringiensis, B. megaterium, B. cereus, B. natto,* and, *B. acidocaldarius*.

Especially the acid α-amylase activity of the invention would be valuable for these strains as lactic acid bacteria are widely used for preservation of many starch containing raw materials (e.g., cereal grains, edible roots and crop residues).

*Lactobacillus* carrying and expressing the acid α-amylase gene of the invention, or a fragment, variant or functional derivative thereof, would be also extremely useful in food and fodder preparation. Such expression may occur under the control of the homologous (native) regulatory regions or under the control of heterologous promoter and/or signal sequences, especially those of *Lactococcus* or *Lactobacillus* origin.

Additionally, the *Lactobacillus* hosts of the invention, carrying the acid amylase of the invention are useful for the preservation of feed and fodder material rich in starch, e.g., cereal grains (siliaged with crimping), corn (maize), maize cobs and alfalfa.

Another specific application of these strains is their use in starter feeds for calves, in milk replacers for calves younger than four weeks (to increase digestibility of starch and for probiotic action) and in starter feed for piglets.

Furthermore, these strains can be effectively used for preservation of material containing cereal and slaughter scraps for fodder use.

The examples below are for illustrative purposes only and are not deemed to limit the scope of the invention.

## EXAMPLES

### Example 1

### Isolation of the Acid α-Amylase Enzyme from Bacillus acidocaldarius

*Bacillus acidocaldarius* (*B. ac.*) strain ATCC 27009 was grown on agar plates containing (per liter): 1 g yeast extract, 0.2 g $(NH_4)_2SO_4$, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.25 g $CaCl_2 \cdot 2H_2O$, 0.6 g $KH_2PO_4$ and 2.5 g maltose, pH 4.5. To isolate the α-amylase enzyme, the entire surface of the agar plate was inoculated with *B. ac.* cells and the plates were incubated 3-5 days at 55°C. The cultures were then suspended in 0.5 M sodium acetate, pH 5.0, and the cells were removed by centrifugation at 8000 g for 20 min. The supernatant was centrifuged again at 40,000 g for 30 min and proteins in 50 ml of the cleared growth medium were precipitated at 0°C by slow addition of $(NH_4)_2SO_4$ to a final concentration of 70% (w/v). After allowing the precipitate of the proteins to form for 30 min at 0°C, the precipitated proteins were collected by centrifugation at 10,000 g for 20 min. The pellet was dissolved in 3 ml of 20 mM BisTris, pH 5.8, and applied onto a Bio-Gel P-200 (Sigma, Richmond, USA) column (1.5 x 45 cm). The elution was carried out in the same buffer. Rapid screening for amylase activity in the eluted fraction was performed using plate assay. The assay plates contained 1.5% agar, 0.2% starch, 20 mM $CaCl_2$, 50 mM sodium acetate, pH 5.0.

Samples were applied into wells made in agar and, after incubation for 2-10 hours at 55°C, the enzyme activity was detected by spreading 0.01 M $I_2$/0.01M KI solution onto the plates. The enzyme activity can be detected as a halo around the agar well. For quantitative assay of the acid α-amylase activity, the Phadebas® amylase test (Pharmacia) was used.

The amylase containing fractions from Bio-Gel P-200 were pooled and concentrated by ultrafiltration in a Novacell™-Omegacell™ apparatus (Filtron, Northborough, Massachusetts, USA). The concentrate was rechromatographed in a Superose 12 HR 10/30 column (Pharmacia, Uppsala, Sweden) in 20 mM Bis-Tris, pH 5.8. The amylase containing fractions were concentrated as above by ultrafiltration and stored at -20°C.

To detect the amylase activity in SDS-polyacrylamide gels, zymography was used (Harris, H. *et al., Handbook of Enzyme Electrophoresis in Human Genetics*, North-Holland, Amsterdam, and, Methods in Enzymology 22:578 (1971). After SDS electrophoresis, the gel was washed for 15 min in 50 mM sodium acetate, pH 5.0 and put onto an amylase assay plate. After incubation at 55°C for 10 hours the enzyme activity was detected as above. When the amylase(s) produced by *B. ac.* strain ATCC 27009, isolated and partially purified as described above, were analyzed by SDS-PAGE and zymography, two forms with relative molecular weights of 90,000 and 160,000 were detected (Figure 1). Similar results were obtained when the SDS-PAGE preceding zymography was performed under nonreducing conditions. This suggests that the 160 kd form was not a disulfide linked dimer of the 90 kd form. Under

the conditions in zymography the 160 kd form had apparently much higher specific activity (Figure 1).

To assay the amylase activity after SDS-PAGE by zymography, the SDS-gel was first incubated in 20 mM $CaCl_2$, 50 mM sodium acetate, pH 5.0 for 15 min at room temperature. AFter this incubation, the SDS-gel was placed on an agarose gel (3mm think, 1.0% agar, 0.2% starch, 20 mM $CaCl_2$, 50 mM sodium acetate, pH 5.0). The gels were then incubated at 55 oC for 10 hours. To reveal the amylase activity, the agarose gel was stained by spreading 0.01M $I_2$/0.01M KI solution onto the gel. A clear halo in the dark background indicated the amylase activity.

In chromatofocusing (Huchthens, T.W. in:*Protein Purification*, J.C. Janson *et al.*, eds., V.H.C. Publishers Inc., New York (1989)) and isoelectric focusing (der Lan *et al.*, in:*Gel Electrophoresis of Proteins*, B.D. Hames *et al.*,eds., IRL Press, Oxford) followed by zymography both the 160 kd and 90 kd polypeptides migrated as a broad peak and band, respectively, at pI 4.8. When the pH optimum of the enzyme(s) was measured, a symmetric curve with a relatively sharp pH optimum at pH 5 was obtained. The enzyme activity was about four fold at 60°C to that at 37°C (Figure 2). The above data suggests that the 90 kd polypeptide might be a degradation product of the 160 kd protein.

Example 2

N-Terminus Sequencing of the Acid α-Amylase Gene

To sequence the acid α-amylase enzyme, SDS-PAGE was performed to separate the 90 kd and 160 kd polypeptides from the partially purified culture supernatant. After SDS-PAGE the gel was treated with 1 M KCl to visualize protein bands and the band corresponding the enzyme activity was excised. From the gel the protein was electroeluted using an ISCO model 1750 electrophoretic concentrator as described in *Advanced Methods in Protein Microsequence Analysis*, Wittmann-Liebold, B., Salnikow, J. and Erdman, V.A., eds., pp. 194-206, 1986, Springer-Verlag Berlin Heidelberg. Only the 90 kd polypeptide could be purified in an amount required for $NH_2$-terminal sequencing. For $NH_2$-terminal amino acid sequence analysis acid amylase was degraded in a gas/pulsed liquid sequencer (*J. Prot. Chem. 7:242-273 (1988)*) after electrophoretic transfer onto a polyvinylidene difluoride membrane Speicher, D.W. in:*Techniques in Protein Chemistry*, T.E.Hugli, ed., Academic Press., San Diego, CA (1989). When the 90 kd protein electrotransferred onto a polyvinyl difluoride membrane was degraded, a single amino acid sequence of $NH_2$-Asp-Ile-Asn-Asp-Tyr was obtained.

For the cloning of the corresponding acid α-amylase gene, eight lysylendopeptidase-cleaved peptides of the 90 kd protein were purified as follows: After electroelution of the 90 kd protein from the gel, the eluate was freeze dried and the solid material was redissolved in 50 µl of 50 mM Tris/HCl, pH 9.0. 150 ng of lysylendopeptidase (Wako,

Dallas, USA) was added and the mixture was incubated at 30°C for 18 hours. The resulting peptides were separated by reverse phase chromatography on a Vydack 218 TPB5 (0.46 x 15 cm) column connected to a Varian 5000 liquid chromatograph. The peptides were eluted using a linear gradient of acetonitrile (0-60% in 90 min) in 0.1% trifluoroacetic acid. The peptides were sequenced after application on polybrene (2mg) pretreated glass fiber filters. The N-terminal sequence of the eight purified peptides derived from 90 kd protein are shown in Table 1. Three of the eight N-terminal sequences (Table 1 no. 2, 3 and 4) were used to synthetize oligonucleotides. The sequences of these oligonucleotides, used for screening of the *B. ac.*-gene library, are shown in Table 2.

TABLE 1

N-terminal sequences of the lysylendopeptidase cleaved peptides of the 90 kd protein.

| Peptide | Sequence |
|---|---|
| 1 | tyr-asp-thr-ala-asp-tyr-phe-lys [SEQ ID No. 6] |
| 2 | leu-val-ser-asp-phe-gln-phe-ser-phe-pro-x-asp-pro-thr-ile-phe-tyr [SEQ ID No. 7] |
| 3 | ile-asp-pro-gly-phe-gly-thr-gln-gln-asp-x-leu-asn-leu-val-gln-ala [SEQ ID No. 8] |
| 4 | his-ala-val-ile-tyr-glu-ile-met-pro-asp-pro-phe-tyr-asn-gly-asn-ile-ala [SEQ ID No. 9] |
| 5 | leu-asp-tyr-leu-lys [SEQ ID No. 10] |
| 6 | ser-leu-gly-val-asn-thr-leu-tyr-leu-met-pro-val-phe-glu-ala[SEQ ID No. 11] |
| 7 | phe-gly-asn-phe-his-ser-asn-gly [SEQ ID No. 12] |
| 8 | gly-ile-tyr-val-gly-ala [SEQ ID No. 13] |

Table 2

Oligonucleotides used for screening of the B.ac. - α-amylase gene. Oligos 370, 371 and 372 correspond to peptides 2, 3 and 4, respectively. "I" is inosine.

| Oligonucleotide | Sequence |
|---|---|
| 370 | 5' TA GAA GAT IGT IGG GTC III IGG GAA IGA GAA-CTG GAA GTC IGA IAC 3' [SEQ ID No. 3] |
| 371 | 5' CA GTC CTG CTG IGT ICC GAA ICC IGG GTC GAT 3' [SEQ ID No. 4] |
| 372 | 5' GC IAT GTT ICC GTT ITA GAA IGG GTC IGG CAT IAT CTC ITA IAT IAC IGC GTG 3' [SEQ ID No. 5] |

## Example 3

## Cloning of the *B. acidocaldarius* Acid α-Amylase Gene

ATCC 27009 *B. ac.* cells were grown in a liquid medium containing [per liter] 1g yeast extract, 0.2g $(NH_4)_2SO_4$, 0.5 g $MgSO_4.7H_2O$, 0.25 g $CaCl_2$, 0.6g $KH_2PO_4$ and 1g glucose, pH 4.5 at 55°C until late logarithmic phase. The cells were collected and the chromosomal DNA was isolated according to Marmur, *J. Mol. Biol.* 3: 208-218 (1986)), except that the DNA was dissolved in saline after first ethanol precipitation. The chromosomal DNA was partially digested with *Hae*III and after agarose gel electrophoresis a fragment population of 4-6 kb was isolated. The DNA fragments were cloned in *E. coli* using λ-gt10 as the vector (Amersham) according to manufacturers recommendation. The *B. ac.* gene bank in *E. coli* was screened by plaque hybridization using oligonucleotides 370 [SEQ ID No. 3]; 371 [SEQ ID No. 4] and 372 [SEQ ID No. 5] (see table 2) as

radioactive probes. After hybridization a positive clone, with an 4.2 kb insert, was found. The insert was subcloned in M13 vectors and pBR322 and sequenced either by M13 method or using direct plasmid sequencing method.

The DNA sequence revealed an open reading frame of 3532 nucleotides starting from the 5' end of the 4.2 kb insert (see Figure 3 and [SEQ ID. No. 1]). Within the open reading frame, starting from the codon 342, a deduced amino acid sequence DINDY (see Figure 3) can be found. This sequence corresponds to the N-terminal amino acid sequence of the 90 kd protein. The open reading frame downstream of codon 342 codes for a 92 kd protein which is in good agreement with the purified $M_r$ 90 kd band. Codon 342 is preceded by an open reading frame at the 5' end of the 4.2 kb clone, and there are no sequences resembling either promoter or SD-regions in the near vicinity of the N-terminal codon (codon 342). This strongly suggests that the 90 kd protein indeed is a breakdown product of a larger polypeptide and the regulatory regions of this polypeptide are outside of the 4.2 kb insert. Therefore, the *B.ac.* gene bank was rescreened using a 5' fragment of the 4.2 kb clone as a probe. This resulted in a new 5.1 kb clone which partially overlapped 5' end of 4.2 kb fragment. The organization of these two clones was confirmed by Southern blotting from the *B.ac.*-genome and the joint region was sequenced also directly from a relevant chromosomal PCR-fragment. The 5.1 kb clone was sequenced by M13 method.

The DNA sequence upstream of the 4.2 kb fragment revealed that open reading frame of the 4.2 kb fragment continued additional 400 codons. The open reading was preceded by a typical promoter - SD-region, and the deduced N-terminal amino acid sequence showed typical features of the bacterid signal peptide (see Figure 3). The combined

open reading frame codes for a protein of 150 kd which is in good agreement with the 160 kd band shown in the SDS-PAGE - zymography. The deduced amino acid sequence revealed sequences that are typical to other amylotic enzymes and surprisingly, a C-terminus that resembles a hydrophobic anchor sequence.

To combine the two parts of the amylase gene the relevant DNA fragments of the 5.1 kb and 4.2 kb clones have been ligated in to a pHP13 plasmid vector (*Mol. Gen. Genetics 209*:335-342 (1987)) and transformed in *Bacillus subtilis*. The DNA analysis of the acid amylase/pHP13 hybrid plasmid confirmed the intact acid amylase gene was stably maintained in *B. subtilis*. This hybrid plasmid was designated pKTH2034.

## Example 4

## Expression of the Acid α-Amylase Gene in *Lactobacillus*

Using the clone described herein, the *B.ac.* α-amylase gene is expressed in *Lactobacillus* host cells. The method is described in detail in Examples V, VI, VII and X, of the Applicants' copending Application Serial No. 377,450, filed July 10, 1989, the specification of which is incorporated herein by reference in its entirety as if at forth in full.

Briefly, plasmids as exemplified by pKTH 1797, pKTH 1798, pKTH 1799, pKTH 1801, pKTH 1805, pKTH 1806, pKTH 1807 and pKTH 1809 contain promoter and secretion promoting signals. The promoter and secretion promoting sequences of these plasmids can direct the expression of heterologous gene and secretion of gene product Gram positive host cells such as *Bacillus*, *Lactobacillus* and *Lactococcus*.

To obtain expression of the acid α-amylase gene and secretion or export of the gene product, a host cell such as *Lactobacillus* is transformed with a plasmid comprising

the promoter and secretion promoting sequences described above and the gene coding for the mature acid amylase protein described herein. In addition of using a plasmid that is able to replicate in *Lactobacillus* host, a nonreplicative plasmid (e.g., pBR322 or pE194), bearing random fragments of *Lactobacillus* chromosome, can be used. Transformation with nonreplicative plasmid results in integration of the acid amylase gene in the *Lactobacillus* chromosome.

The host containing the amylase gene is cultured in a suitable medium allowing expression of the protein and the protein is recovered from the cell wall layer or culture medium or the host cell, producing the α-amylase activity, is directly used in the required process. The *Lactobacillus* host of the invention, modified to express the acid=α-amylase gene, may be utilized for the processing of food and feed as in the following examples.

## Example 5
## Processing of Forage

Forage, for instance grass, is harvested by either a frail harvester or a precision chopper attached to a wagon of a tractor. Preservation solution is added to the chopped forage mass during the harvesting in the chopper/harvester using equipment designed for this purpose. Th preservation solution is prepared by adding an inoculant (a *Lactobacillus* strain as described above, capable of producing the acid amylase of the invention), to a bottle or tank which contains half of its volume as lukewarm water. This water should not be so hot as to kill the bacteria. After shaking to evenly distribute the inoculant, the rest of the volume of the bottle is filled

by lukewarm water. Usually farmers use bottles of 30 liters or tanks of 200 liters. The inoculant can be either in liquid form or as a dried powder of the acid amylase-expressing hosts of the invention. The inoculant may be prepared in fermentors using whey-based medium and the culture is grown to a cell density of approximately $10^9$ cells/ml. After the culture has reached this density, the cells are collected by centrifugation and lyophilized to a final density of $10^{11}$-$10^{12}$ cells/g of dry weight. Twenty grams of this lyophilized material or equal amount of cells in a liquid concentrate is mixed with 30 liters of water as described above. Five liters of this diluted solution is applied per ton of forage.

Alternatively, the prepared preservation solution can also be added to the forage mass after the wagon has been emptied in the silo in which fodder is ensiled.

## Example 6
## Animal Feed Additive

A concentrate mixture (powdered, granulated or pelleted) is manufactured using standard manufacturing systems. Each component in the mixture is separately weighed and then mixed together before pelleting, for instance. An inoculant (as described in Example 5) can be added or mixed into the mixture as one of the components. The concentrate, including the inoculant, can be fed to domestic animals, such as, for example, cows, calves, pigs, piglets, etc.

The inoculant can also be fed separately as a specially formulated powder, tablet or a liquid product. The inoculant is generally included in a carrier ingredient, for instance, skim milk powder. These products can be fed to domestic animals directly, as as a component together with the other feeds or in the drinking water of the animal.

For each purpose, an inoculant is manuactured so that it contains $10^{11}$-$10^{12}$ cfu/g dry inoculant powder. The inoculant may be used in any concentration which is necessary to achieve the desired effect. Especially, when used as an additive to the animals' feed, the inoculant is used at 0.01-0.02% in the concentrate mixture.

The advantage of feeding the inoculant of the invention to domesticated farm animals such as the above is that by letting the hosts of the invention colonize the intestional tract of the animal, the animal is better able to digest the silage or fodder that it is fed, as the acid stable amylase secreted by such hosts would better degrade the ingested foodstuff.

All references cited herein are incorporated herein by reference. While this invention has been described in detail and with reference to specific embodiments thereof, it will be apparent to one skilled in the art that various

changes and modifications could be made therein without departing from the spirit and scope thereof.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Teija Koivula,
Mervi Sibakov, and
Ilkka Palva

(ii) TITLE OF INVENTION: Acid α-Amylase
(iii) NUMBER OF SEQUENCES: 13

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Sterne, Kessler, Goldstein & Fox
(B) STREET: 1225 Connecticut Avenue, N.W.,
Suite 300
(C) CITY: Washington
(D) STATE: District of Columbia
(E) COUNTRY: United States of America
(F) ZIP: 20036

(v) COMPUTER READABLE FORM;
(A) MEDIUM TYPE: Floppy Disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: Ascii

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Michele Ann Cimbala
(B) REGISTRATION NUMBER: 33,851
(C) REFERENCE/DOCKET NUMBER: 1155.0090000

(viii) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (202) 833-7533

(2) INFORMATION FOR SEQ ID NO :1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4140 bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double Stranded
(D) TOPOLOGY: Linear
(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CATGTGGGTC GCCTCCTCTG CGATGGGACA TCTCACTTAC ATTTTAACAG TGGAAGGGTG 60

GGGAAGGGAA CAAGGGTCTT GAGAGAAAAG AAAGGATTCG GACAATGAAC GATGCAAGAG 120

CACCATTGAC GTCGGGAAGG CTGAGCGTAT AATCCAGGAC GACAACGTTA GTGCAATCGT 180

TTGCACAAGA GGAGGTTGAT GCACGTGAAG TGGACGGGTG CCGCGAGCCT TGCGACGACC 240

ATCGTCATGT CTGCAGCGCT TCCGTTCGCG GCGTTTGAAA ACACCCGTTC CGTCGGCACG 300

GAACGCGTTC TGGCGGCGAC CGCGGCGGAT AGTACTAACG CCTCGCCCCC GACCGGGGCG 360

TCCGCAGGCG GTGGCACAGC AGGAGAGACC TACTCGAATG CCGTGCAGAT TGTGAGCGGC 420

AATGGCTTTC GCGCTGGAAA CGGGGCCACG GTCGTCCTCG CTGTGAATGG ACTGAACTTG 480

AACCCCTCGA GTTTCAAGGT TGTGGTATCG AATAGCCTCA ACGGCGTGAT CGACGTCACA 540

AGTGATTCCG TGCTAACGGG GCCGAATTCG ATTGCCTTGC ACCTTCCTGC CGGGGATGAC 600

GGGCTTGGCG CCGGTACGTA CACGGTCACC GTCGAAAGTG GAGGTGATTC GGTCTCGACG 660

CCTTCAGGAC AGGGACTTCA GGTCATGCCT TATACGACGG CCGACACCAT TCAGTGGGAT 720

GGAATCTACA CATCCGATGG CGCGATGTAC GTGTCAGATC CGAATCCCTC CCCTGGACAA 780

GAGGTGACGA TCAGCCTTCG AGCGTACAGC GGGAACCTGA CGAAGGTGAT CCTGAATTGC 840

TGGGACACCG CACAAAACAA GGGTTTCCAG GTCGAGATGT CTCCAGGCAG AACGTTTGGA 900

CCGTATCAGC TCTGGTCTGC CACGATCCCC GCTTCAAACG GCGGAACGAT CTATTATCGC 960

TTTGACATAT ACGATGGCAC CAGTTTTGCG TGTCTCTCAG GTGACGGACT GCACACGTCT 1020

GACGACATCA ACAACAATTT CCCGTTGCCC GTGGGGACGG TCACGCTTTC GACACTTCAG 1080

GCGAATCCCG GTGATACGGT GACGGTCTCC GACCCTGTAG GTGACTTCGC CGGAAGCCAG 1140

GATCAACCCA ATCACACGGT GATACGGTTT GTCAACTCGT CGGGCGAAAC GGCCGCCACA 1200

GTCAATGGGA CGAACGCGAG CTGGAACAGC GTGCAGTTCA CAGTCCCACA GAGCCTTCCA 1260

AACGGCTTGT ATCGCGTCGA GATCGACACG GTCGCCAAGG ACGCGGATGG GGTGGTCAAT 1320

GTCGAATTGG ACAGGAGTGC GGAGCTTATT GTAGGGCCTC TGCCCGCGTG GATGCAAGCG 1380

TATGCACATG ATTCGTTTCA GGCGTTCTAC CGATCGCCTT TCGGAGCCGT GTCCACAGGA 1440

ACCCCCATCA CGCTTCGCCT GCGGGCTCCG CTCAGCGTGA AGAGTGCGAC GCTTCGCCTC 1500

TGGGGGGCAG CGGATCAGTC AGGCGAGATC GACCTGCCGA TGCAGAAGCT CCAAATGTCG 1560

GGAGACGAGT TGGCGCAACA AACCGGCGTG CAGGACATCA ACGACTACAC GTGGTGGACG 1620

GTGACCATCC CTGCGGCGGA TGTGACCACA CCCGGGACGA TGTGGTATCA GTTCGTGACG 1680

GAGACGGACA CTGGCCAGGT GGTCTACTAC GATGACAATG GAGCTCAGCT TGAAGGGCCT 1740

GGCCAGGTTG GGTTGTCTTC CGACGGACCG AGCTACCAGA TCAGCGTATA CGAACGGGGA 1800

TTTCAGACGC CAGATTGGCT GAAGCACGCC GTGATCTACG AAATCATGCC GGATCGGTTC 1860

TACAATGGCA ATATCGCCAC GGAGGAGAAT CCGAATACGC AAAAGGGGAT TTATGTAGGG 1920

GCCGATGGAA CGGAGTCATT AGGCCCCATC CAGTTCCACG AGAACTGGGA CTCGCCGCCC 1980

TATGATCCGA ATATTCCTCC GTTATCTGAT CCCAAAATTG CCAGTCTGCG AGGCAATGGC 2040

CAATGGAACA TTGACTTTTT CGGAGGTGAT TTGCAGGGCA TCGAGGATAA GCTGGACTAC 2100

CTGAAGAGCC TTGGAGTCAA TACGCTGTAT CTGATGCCCG TCTTTGAGGC GGAATCCAAT 2160

CACAAATATG ACACAGCCGA CTATTTCAAG ATTGACCCTG GATTTGGAAC GCAGCAGGAC 2220

TGGCTGAATC TCGTACAGGC TGCGCACGCG AAGGGGTTCC ATATCATTCT CGACGGGGTG 2280

TTCGAAGATA CCGGAAGTGA CAGCGTATAT TTCAACAAGT TCGGGAACTT CCACTCCAAC 2340

GGTGCGTGGC AGGCGTACCT GAAGAACCAG CCGTCGCTGT CGCCCTACTA CTCGTGGTAC 2400

GTGTGGACAG GGAACACCTC AAACCCATAC GATTCGTGGT TTCAGATCGA CACGCTGCCA 2460

CTTACGGACA CGTCGAACCC CGCCTATCAG CGATTCGTGT ATGGGAGCGA CAACTCAGTC 2520

GCGCGTGTGT GGATCCGGGA AGGTGCGGAC GGATGGCGCT TGGACTCGGC CGACAACGGG 2580

AATTTCAACA CGGCATGGTG GGGTGGCTTT CGGCAGGCCG TGAAATCGAT CGATCCCAAC 2640

GCAGCGATCA TCGGCGAGAT CTGGGACAAT GCGACGAATG ACAATGGAAC GGATTGGTTG 2700

ACGGGATCGA CCTTCGACAG TGTAATGAAC TACCAGTTCC GGAACGCCGT GATCGACTTC 2760

TTCCGCGGCA CGTACAACGA CGGAAACGTG CAGCACCACG CCGTCGACGC TGCGGGATTC 2820

AACCAGGAAC TGATGCGCCT GTACAGTGAA TATCCTCTGC AGTCGTTCTA CTCGATGATG 2880

AACCTTGTCG ATTCGCAAGA CACCATGCGG ATCCTGACCA TCTTGGAGAA CGCGCCGCAG 2940

CCAGGCGATC TATCCGCGCT CCAGCAGGAT GAGTACAGGC CGTCTCCTGC GGCTGAACAG 3000

TTGGGGATCG AGAGGCTGAA GCTTGTATCG GACTTTCAAT TCAGCTTCCC GGGCGATCCG 3060

ACCATCTTCT ACGGCGACGA GGCAGGGCTC ACTGGTTATT CGGATCCCCT CAATCGTCGG 3120

ACCTATCCGT GGGACAACCA GAATCTCGAT CTCCTGAACC ACTACCGCAA GCTCGGGGCC 3180

ATTCGAAACG CCAATCCTGT GCTTCAGACG GGGGATTTCA CGCCGTTGTA CGCACAGGGC 3240

ATGGTGTACG CATTTGCAAG GACCATTCGG AATGGGCGAG ATGTCTTCGG TGTGCCAGCG 3300

GAGGATGCCA CGGCCATTGT GGCGATCAAC AATCAGAACC AAGCTATCAC CGTGACCATT 3360

CCGACGGATG GGACGGTTGC GGACGGGTCC ACGATGCTCG ATGAACTGAA CAACCAGTGG 3420

TACAAGGTGC AGAATGGTGG CATTACACTC ACGCTGCAAT CGTATCAAGG TGCCATTTTG 3480

GTGACGCCGA GCGACGCGCC GATGGCTTAT CTGCAAGAGG AGGATTCTCA GAACGAGATT 3540

GCGTGGACGC CTGTGCAAGG TGCCATCGGT TATCGCGTCT GGAGACAGAA TCCGAATGGA 3600

CAATGGGTGC CCTTTGGACC TGTGCTTCCT GCCACGGACT TGAGTGTCAC GGTGGAACGC 3660

GATGCATATG CGCAAACGTT TGCTGTACAA GCGCTGTTTT CGGCGTCTGA TCACGCCCAG 3720

TCTCCGGTGT CGGCACCTAA GACGGTATCG CTTCCCGTCG ATGTGCCCGC GGTACGCCTG 3780

AGTCAGCCGA TCGTTAGTGG TCGTGTGGTT GGAGATCGTG CGATGGTCTC GATCACGCCG 3840

GTTTCAGGCG CGACGCAGTA TGTGATCTAC CAGAGACAGG GCGACGGATC GTATGCTCCG 3900

GTCGCGACGG TCTCCACAAG TGGCGATTCC GCAGCTATAG GGGAAGTTCC TGCGCAAGGT 3960

CCGGCCAACT CGCCTCACGC GACGATTCGC GTGACAGTGC CCGTACCTGC AGGTTTCTCG 4020

TCGGTGACCT ACCGCGTGGC TGCGCAAAAC GAAGATGGGC AAGCTGTGAC CAATCCATTG 4080

ACCCTATCGC TCTCGAAAAA GTGATGCCTC GCGAAAAGGG CTATCGGAAT TTTTTCGAGA 4140

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1341 amino acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met His Val Lys Trp Thr Gly Ala Ala Ser Leu Ala Thr Thr Ile Val
 1               5                  10                  15

Met Ser Ala Ala Leu Pro Phe Ala Ala Phe Glu Asn Thr Arg Ser Val
            20                  25                  30

Gly Thr Glu Arg Val Leu Ala Ala Thr Ala Ala Asp Ser Thr Asn Ala
        35                  40                  45

Ser Pro Pro Thr Gly Ala Ser Ala Gly Gly Gly Thr Ala Gly Glu Thr
    50                  55                  60

Tyr Ser Asn Ala Val Gln Ile Val Ser Gly Asn Gly Phe Arg Ala Gly
65                  70                  75                  80

Asn Gly Ala Thr Val Val Leu Ala Val Asn Gly Leu Asn Leu Asn Pro
                85                  90                  95

Ser Ser Phe Lys Val Val Val Ser Asn Ser Leu Asn Gly Val Ile Asp
            100                 105                 110

Val Thr Ser Asp Ser Val Leu Thr Gly Pro Asn Ser Ile Ala Leu His
        115                 120                 125

Leu Pro Ala Gly Asp Asp Gly Leu Gly Ala Gly Thr Tyr Thr Val Thr
    130                 135                 140

Val Glu Ser Gly Gly Asp Ser Val Ser Thr Pro Ser Gly Gln Gly Leu
145                 150                 155                 160

Gln Val Met Pro Tyr Thr Thr Ala Asp Thr Ile Gln Trp Asp Gly Ile
                165                 170                 175
```

```
Tyr Thr Ser Asp Gly Ala Met Tyr Val Ser Asp Pro Asn Pro Ser Pro
            180                 185                 190

Gly Gln Glu Val Thr Ile Ser Leu Arg Ala Tyr Ser Gly Asn Leu Thr
        195                 200                 205

Lys Val Ile Leu Asn Cys Trp Asp Thr Ala Gln Asn Lys Gly Phe Gln
    210                 215                 220

Val Glu Met Ser Pro Gly Arg Thr Phe Gly Pro Tyr Gln Leu Trp Ser
225                 230                 235                 240

Ala Thr Ile Pro Ala Ser Asn Gly Gly Thr Ile Tyr Tyr Arg Phe Asp
                245                 250                 255

Ile Tyr Asp Gly Thr Ser Phe Ala Cys Leu Ser Gly Asp Gly Leu His
            260                 265                 270

Thr Ser Asp Asp Ile Asn Asn Asn Phe Pro Leu Pro Val Gly Thr Val
        275                 280                 285

Thr Leu Ser Thr Leu Gln Ala Asn Pro Gly Asp Thr Val Thr Val Ser
    290                 295                 300

Asp Pro Val Gly Asp Phe Ala Gly Ser Gln Asp Gln Pro Asn His Thr
305                 310                 315                 320

Val Ile Arg Phe Val Asn Ser Ser Gly Glu Thr Ala Ala Thr Val Asn
                325                 330                 335

Gly Thr Asn Ala Ser Trp Asn Ser Val Gln Phe Thr Val Pro Gln Ser
            340                 345                 350

Leu Pro Asn Gly Leu Tyr Arg Val Glu Ile Asp Thr Val Ala Lys Asp
        355                 360                 365

Ala Asp Gly Val Val Asn Val Glu Leu Asp Arg Ser Ala Glu Leu Ile
    370                 375                 380

Val Gly Pro Leu Pro Ala Trp Met Gln Ala Tyr Ala His Asp Ser Phe
385                 390                 395                 400

Gln Ala Phe Tyr Arg Ser Pro Phe Gly Ala Val Ser Thr Gly Thr Pro
                405                 410                 415

Ile Thr Leu Arg Leu Arg Ala Pro Leu Ser Val Lys Ser Ala Thr Leu
            420                 425                 430

Arg Leu Trp Gly Ala Ala Asp Gln Ser Gly Glu Ile Asp Leu Pro Met
        435                 440                 445

Gln Lys Leu Gln Met Ser Gly Asp Glu Leu Ala Gln Gln Thr Gly Val
    450                 455                 460

Gln Asp Ile Asn Asp Tyr Thr Trp Trp Thr Val Thr Ile Pro Ala Ala
465                 470                 475                 480

Asp Val Thr Thr Pro Gly Thr Met Trp Tyr Gln Phe Val Thr Glu Thr
                485                 490                 495
```

```
Asp Thr Gly Gln Val Val Tyr Tyr Asp Asp Asn Gly Ala Gln Leu Glu
            500                 505                 510

Gly Pro Gly Gln Val Gly Leu Ser Ser Asp Gly Pro Ser Tyr Gln Ile
        515                 520                 525

Ser Val Tyr Glu Arg Gly Phe Gln Thr Pro Asp Trp Leu Lys His Ala
    530                 535                 540

Val Ile Tyr Glu Ile Met Pro Asp Arg Phe Tyr Asn Gly Asn Ile Ala
545                 550                 555                 560

Thr Glu Glu Asn Pro Asn Thr Gln Lys Gly Ile Tyr Val Gly Ala Asp
                565                 570                 575

Gly Thr Glu Ser Leu Gly Pro Ile Gln Phe His Glu Asn Trp Asp Ser
            580                 585                 590

Pro Pro Tyr Asp Pro Asn Ile Pro Pro Leu Ser Asp Pro Lys Ile Ala
        595                 600                 605

Ser Leu Arg Gly Asn Gly Gln Trp Asn Ile Asp Phe Phe Gly Gly Asp
    610                 615                 620

Leu Gln Gly Ile Glu Asp Lys Leu Asp Tyr Leu Lys Ser Leu Gly Val
625                 630                 635                 640

Asn Thr Leu Tyr Leu Met Pro Val Phe Glu Ala Glu Ser Asn His Lys
                645                 650                 655

Tyr Asp Thr Ala Asp Tyr Phe Lys Ile Asp Pro Gly Phe Gly Thr Gln
            660                 665                 670

Gln Asp Trp Leu Asn Leu Val Gln Ala Ala His Ala Lys Gly Phe His
        675                 680                 685

Ile Ile Leu Asp Gly Val Phe Glu Asp Thr Gly Ser Asp Ser Val Tyr
    690                 695                 700

Phe Asn Lys Phe Gly Asn Phe His Ser Asn Gly Ala Trp Gln Ala Tyr
705                 710                 715                 720

Leu Lys Asn Gln Pro Ser Leu Ser Pro Tyr Tyr Ser Trp Tyr Val Trp
                725                 730                 735

Thr Gly Asn Thr Ser Asn Pro Tyr Asp Ser Trp Phe Gln Ile Asp Thr
            740                 745                 750

Leu Pro Leu Thr Asp Thr Ser Asn Pro Ala Tyr Gln Arg Phe Val Tyr
        755                 760                 765

Gly Ser Asp Asn Ser Val Ala Arg Val Trp Ile Arg Glu Gly Ala Asp
    770                 775                 780

Gly Trp Arg Leu Asp Ser Ala Asp Asn Gly Asn Phe Asn Thr Ala Trp
785                 790                 795                 800

Trp Gly Gly Phe Arg Gln Ala Val Lys Ser Ile Asp Pro Asn Ala Ala
                805                 810                 815
```

```
Ile Ile Gly Glu Ile Trp Asp Asn Ala Thr Asn Asp Asn Gly Thr Asp
            820                 825                 830

Trp Leu Thr Gly Ser Thr Phe Asp Ser Val Met Asn Tyr Gln Phe Arg
        835                 840                 845

Asn Ala Val Ile Asp Phe Phe Arg Gly Thr Tyr Asn Asp Gly Asn Val
    850                 855                 860

Gln His His Ala Val Asp Ala Ala Gly Phe Asn Gln Glu Leu Met Arg
865                 870                 875                 880

Leu Tyr Ser Glu Tyr Pro Leu Gln Ser Phe Tyr Ser Met Met Asn Leu
                885                 890                 895

Val Asp Ser Gln Asp Thr Met Arg Ile Leu Thr Ile Leu Glu Asn Ala
            900                 905                 910

Pro Gln Pro Gly Asp Leu Ser Ala Leu Gln Gln Asp Glu Tyr Arg Pro
        915                 920                 925

Ser Pro Ala Ala Glu Gln Leu Gly Ile Glu Arg Leu Lys Leu Val Ser
    930                 935                 940

Asp Phe Gln Phe Ser Phe Pro Gly Asp Pro Thr Ile Phe Tyr Gly Asp
945                 950                 955                 960

Glu Ala Gly Leu Thr Gly Tyr Ser Asp Pro Leu Asn Arg Arg Thr Tyr
                965                 970                 975

Pro Trp Asp Asn Gln Asn Leu Asp Leu Leu Asn His Tyr Arg Lys Leu
            980                 985                 990

Gly Ala Ile Arg Asn Ala Asn Pro Val Leu Gln Thr Gly Asp Phe Thr
        995                1000                1005

Pro Leu Tyr Ala Gln Gly Met Val Tyr Ala Phe Ala Arg Thr Ile Arg
   1010                1015                1020

Asn Gly Arg Asp Val Phe Gly Val Pro Ala Glu Asp Ala Thr Ala Ile
1025                1030                1035                1040

Val Ala Ile Asn Asn Gln Asn Gln Ala Ile Thr Val Thr Ile Pro Thr
               1045                1050                1055

Asp Gly Thr Val Ala Asp Gly Ser Thr Met Leu Asp Glu Leu Asn Asn
           1060                1065                1070

Gln Trp Tyr Lys Val Gln Asn Gly Gly Ile Thr Leu Thr Leu Gln Ser
       1075                1080                1085

Tyr Gln Gly Ala Ile Leu Val Thr Pro Ser Asp Ala Pro Met Ala Tyr
   1090                1095                1100

Leu Gln Glu Glu Asp Ser Gln Asn Glu Ile Ala Trp Thr Pro Val Gln
1105                1110                1115                1120

Gly Ala Ile Gly Tyr Arg Val Trp Arg Gln Asn Pro Asn Gly Gln Trp
               1125                1130                1135
```

```
Val Pro Phe Gly Pro Val Leu Pro Ala Thr Asp Leu Ser Val Thr Val
          1140                1145                1150

Glu Arg Asp Ala Tyr Ala Gln Thr Phe Ala Val Gln Ala Leu Phe Ser
      1155                1160                1165

Ala Ser Asp His Ala Gln Ser Pro Val Ser Ala Pro Lys Thr Val Ser
  1170                1175                1180

Leu Pro Val Asp Val Pro Ala Val Arg Leu Ser Gln Pro Ile Val Ser
1185                1190                1195                1200

Gly Arg Val Val Gly Asp Arg Ala Met Val Ser Ile Thr Pro Val Ser
              1205                1210                1215

Gly Ala Thr Gln Tyr Val Ile Tyr Gln Arg Gln Gly Asp Gly Ser Tyr
          1220                1225                1230

Ala Pro Val Ala Thr Val Ser Thr Ser Gly Asp Ser Ala Ala Ile Gly
      1235                1240                1245

Glu Val Pro Ala Gln Gly Pro Ala Asn Ser Pro His Ala Thr Ile Arg
  1250                1255                1260

Val Thr Val Pro Val Pro Ala Gly Phe Ser Ser Val Thr Tyr Arg Val
1265                1270                1275                1280

Ala Ala Gln Asn Glu Asp Gly Gln Ala Val Thr Asn Pro Leu Thr Leu
              1285                1290                1295

Ser Leu Ser Lys Lys
          1300
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double Stranded
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TAGAAGATIG TIGGGTCIII IGGGAAIGAG AACTGGAAGT CIGAIAC 47

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double Stranded
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CAGTCCTGCT GIGTICCGAA ICCIGGGTCG AT 32

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double Stranded
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE: Nucleic Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCIATGTTIC CGTTITAGAA IGGGTCIGGCA TIATCTCITA IATIACIGCG TG 52

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 Amino Acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TYR ASP THR ALA ASP TYR PHE LYS
1 5 8

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 Amino Acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

LEU VAL SER ASP PHE GLN PHE SER PHE PRO Xaa ASP PRO THR ILE PHE
1 5 10 15
TYR
17

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 Amino Acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ILE ASP PRO GLY PHE GLY THR GLN GLN ASP Xaa LEU ASN LEU VAL GLN
1 5 10 15

ALA
17

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 Amino Acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

HIS ALA VAL ILE TYR GLU ILE MET PRO ASP PRO PHE TYR ASN GLY ASN
1 5 10 15

ILE ALA
18

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 Amino Acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

LEU ASP TYR LEU LYS
1 5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 Amino Acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

SER LEU GLY VAL ASN THR LEU TYR LEU MET PRO VAL PHE GLU ALA
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 Amino Acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

PHE GLY ASN PHE HIS SER ASN GLY
1 5 8

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 Amino Acids
(B) TYPE: Amino Acid
(C) STRANDEDNESS:
(D) TOPOLOGY: Linear

(ii) MOLECULAR TYPE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GLY ILE TYR VAL GLY ALA
1 5 6

WHAT IS CLAIMED IS:

1. Substantially pure acid α-amylase DNA.

2. The substantially pure acid α-amylase DNA of claim 1, wherein the sequence of said DNA is that of SEQ ID No. 1, or a biologically active fragment, variant, or functional derivative thereof.

3. A recombinant DNA molecule, wherein said recombinant DNA molecule comprises the DNA of claim 2.

4. A cloning vector, wherein said cloning vector comprises the recombinant DNA molecule of claim 3.

5. The cloning vector of claim 4, wherein said recombinant DNA molecule is capable of being expressed.

6. A host cell transformed with the recombinant DNA molecule of claim 3.

7. The host cell of claim 6, wherein said host cell is a bacterial host.

8. The host cell of claim 7, wherein said host cell is a *Lactobacillus* or a *Bacillus*.

9. The host cell of claim 8, wherein said host cell is *Lactobacillus*.

10. The host cell of claim 8, wherein said host cell is *Bacillus*.

11. A method for the production of acid αamylase, wherein said method comprises:

(i) providing a nucleic acid molecule comprising the nucleic acid seq of SEQ ID No. 1, or a fragment, variant or functional derivative thereof;

(ii) transforming a host with said molecule;

(iii) expressing said nucleic acid molecule in said host; and

(iv) isolating the acid α-amylase produced by said expression.

12. The method of claim 11, wherein said host is *Lactobacillus*.

13. The method of claim 12, wherein said host is *Bacillus*.

14. A substantially purified *B. acidocaldarius* acid α-amylase, wherein said acid α-amylase comprises the amino acid sequence of SEQ ID No. 2, or a biologically active fragment, variant or functional derivative therof.

15. The substantially purified acid α-amylase of claim 14, wherein said acid α-amylase is produced by a method comprising:

(i) providing a nucleic acid molecule comprising the nucleic acid seq of SEQ ID No. 1, or a biologically active fragment, variant or functional derivative thereof;

(ii) transforming a host with waid molecule;

(iii) expressing said nucleic acid molecule in said host; and

(iv) isolating the acid α-amylase produced by said expression.

16. A method for preserving feed or fodder, wherein said method comprises the addition of the host of claim 6 to said feed or fodder.

17. The method of claim 16, wherein said host is *Lactobacillus*.

18. The method of claim 16, wherein said feed is cereal grain.

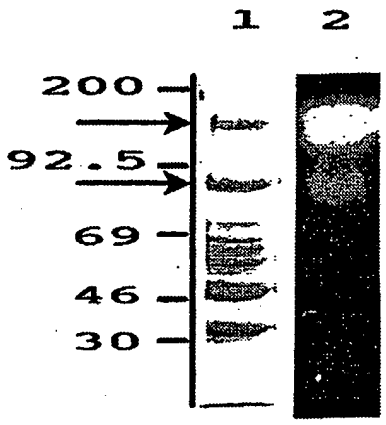


Fig. 1

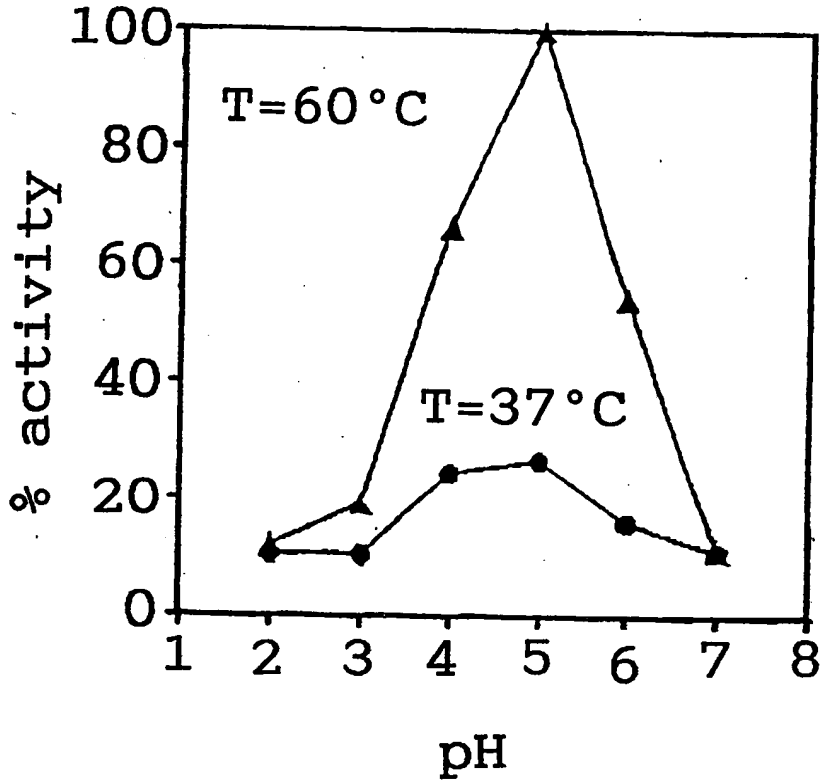


Fig. 2

Fig. 3 - page 1 of 6

B.Acidocaldarius amylase: DNA sequence

```
        10         20         30         40         50         60
         |          |          |          |          |          |
CATGTGGGTCGCCTCCTCTGCGATGGGACATCTCACTTACATTTTAACAGTGGAAGGGTG

        70         80         90        100        110        120
GGGAAGGGAACAAGGGTCTTGAGAGAAAAGAAAGGATTCGGACAATGAACGATGCAAGAG

       130        140        150        160        170        180
CACCATTGACGTCGGGAAGGCTGAGCGTATAATCCAGGACGACAACGTTAGTGCAATCGT

       190        200        210        220        230        240
TTGCACAAGAGGAGGTTGATGCACGTGAAGTGGACGGGTGCCGCGAGCCTTGCGACGACC
                  MetHisValLysTrpThrGlyAlaAlaSerLeuAlaThrThr

       250        260        270        280        290        300
ATCGTCATGTCTGCAGCGCTTCCGTTCGCGGCGTTTGAAAACACCCGTTCCGTCGGCACG
IleValMetSerAlaAlaLeuProPheAlaAlaPheGluAsnThrArgSerValGlyThr

       310        320        330        340        350        360
GAACGCGTTCTGGCGGCGACCGCGGCGGATAGTACTAACGCCTCGCCCCCGACCGGGGCG
GluArgValLeuAlaAlaThrAlaAlaAspSerThrAsnAlaSerProProThrGlyAla

       370        380        390        400        410        420
TCCGCAGGCGGTGGCACAGCAGGAGAGACCTACTCGAATGCCGTGCAGATTGTGAGCGGC
SerAlaGlyGlyGlyThrAlaGlyGluThrTyrSerAsnAlaValGlnIleValSerGly

       430        440        450        460        470        480
AATGGCTTTCGCGCTGGAAACGGGGCCACGGTCGTCCTCGCTGTGAATGGACTGAACTTG
AsnGlyPheArgAlaGlyAsnGlyAlaThrValValLeuAlaValAsnGlyLeuAsnLeu

       490        500        510        520        530        540
AACCCCTCGAGTTTCAAGGTTGTGGTATCGAATAGCCTCAACGGCGTGATCGACGTCACA
AsnProSerSerPheLysValValValSerAsnSerLeuAsnGlyValIleAspValThr

       550        560        570        580        590        600
AGTGATTCCGTGCTAACGGGGCCGAATTCGATTGCCTTGCACCTTCCTGCCGGGGATGAC
SerAspSerValLeuThrGlyProAsnSerIleAlaLeuHisLeuProAlaGlyAspAsp

       610        620        630        640        650        660
GGGCTTGGCGCCGGTACGTACACGGTCACCGTCGAAAGTGGAGGTGATTCGGTCTCGACG
GlyLeuGlyAlaGlyThrTyrThrValThrValGluSerGlyGlyAspSerValSerThr

       670        680        690        700        710        720
CCTTCAGGACAGGGACTTCAGGTCATGCCTTATACGACGGCCGACACCATTCAGTGGGAT
ProSerGlyGlnGlyLeuGlnValMetProTyrThrThrAlaAspThrIleGlnTrpAsp

       730        740        750        760        770        780
GGAATCTACACATCCGATGGCGCGATGTACGTGTCAGATCCGAATCCCTCCCCTGGACAA
GlyIleTyrThrSerAspGlyAlaMetTyrValSerAspProAsnProSerProGlyGln
```

Fig. 3 - page 2 of 6

```
       790        800        810        820        830        840
GAGGTGACGATCAGCCTTCGAGCGTACAGCGGGAACCTGACGAAGGTGATCCTGAATTGC
GluValThrIleSerLeuArgAlaTyrSerGlyAsnLeuThrLysValIleLeuAsnCys

       850        860        870        880        890        900
TGGGACACCGCACAAAACAAGGGTTTCCAGGTCGAGATGTCTCCAGGCAGAACGTTTGGA
TrpAspThrAlaGlnAsnLysGlyPheGlnValGluMetSerProGlyArgThrPheGly

       910        920        930        940        950        960
CCGTATCAGCTCTGGTCTGCCACGATCCCCGCTTCAAACGGCGGAACGATCTATTATCGC
ProTyrGlnLeuTrpSerAlaThrIleProAlaSerAsnGlyGlyThrIleTyrTyrArg

       970        980        990       1000       1010       1020
TTTGACATATACGATGGCACCAGTTTTGCGTGTCTCTCAGGTGACGGACTGCACACGTCT
PheAspIleTyrAspGlyThrSerPheAlaCysLeuSerGlyAspGlyLeuHisThrSer

      1030       1040       1050       1060       1070       1080
GACGACATCAACAACAATTTCCCGTTGCCCGTGGGGACGGTCACGCTTTCGACACTTCAG
AspAspIleAsnAsnAsnPheProLeuProValGlyThrValThrLeuSerThrLeuGln

      1090       1100       1110       1120       1130       1140
GCGAATCCCGGTGATACGGTGACGGTCTCCGACCCTGTAGGTGACTTCGCCGGAAGCCAG
AlaAsnProGlyAspThrValThrValSerAspProValGlyAspPheAlaGlySerGln

      1150       1160       1170       1180       1190       1200
GATCAACCCAATCACACGGTGATACGGTTTGTCAACTCGTCGGGCGAAACGGCCGCCACA
AspGlnProAsnHisThrValIleArgPheValAsnSerSerGlyGluThrAlaAlaThr

      1210       1220       1230       1240       1250       1260
GTCAATGGGACGAACGCGAGCTGGAACAGCGTGCAGTTCACAGTCCCACAGAGCCTTCCA
ValAsnGlyThrAsnAlaSerTrpAsnSerValGlnPheThrValProGlnSerLeuPro

      1270       1280       1290       1300       1310       1320
AACGGCTTGTATCGCGTCGAGATCGACACGGTCGCCAAGGACGCGGATGGGGTGGTCAAT
AsnGlyLeuTyrArgValGluIleAspThrValAlaLysAspAlaAspGlyValValAsn

      1330       1340       1350       1360       1370       1380
GTCGAATTGGACAGGAGTGCGGAGCTTATTGTAGGGCCTCTGCCCGCGTGGATGCAAGCG
ValGluLeuAspArgSerAlaGluLeuIleValGlyProLeuProAlaTrpMetGlnAla

      1390       1400       1410       1420       1430       1440
TATGCACATGATTCGTTTCAGGCGTTCTACCGATCGCCTTTCGGAGCCGTGTCCACAGGA
TyrAlaHisAspSerPheGlnAlaPheTyrArgSerProPheGlyAlaValSerThrGly

      1450       1460       1470       1480       1490       1500
ACCCCCATCACGCTTCGCCTGCGGGCTCCGCTCAGCGTGAAGAGTGCGACGCTTCGCCTC
ThrProIleThrLeuArgLeuArgAlaProLeuSerValLysSerAlaThrLeuArgLeu

      1510       1520       1530       1540       1550       1560
TGGGGGGCAGCGGATCAGTCAGGCGAGATCGACCTGCCGATGCAGAAGCTCCAAATGTCG
TrpGlyAlaAlaAspGlnSerGlyGluIleAspLeuProMetGlnLysLeuGlnMetSer
```

Fig. 3 - page 3 of 6

```
      1570       1580       1590       1600       1610       1620
GGAGACGAGTTGGCGCAACAAACCGGCGTGCAGGACATCAACGACTACACGTGGTGGACG
GlyAspGluLeuAlaGlnGlnThrGlyValGlnAspIleAsnAspTyrThrTrpTrpThr

      1630       1640       1650       1660       1670       1680
GTGACCATCCCTGCGGCGGATGTGACCACACCCGGGACGATGTGGTATCAGTTCGTGACG
ValThrIleProAlaAlaAspValThrThrProGlyThrMetTrpTyrGlnPheValThr

      1690       1700       1710       1720       1730       1740
GAGACGGACACTGGCCAGGTGGTCTACTACGATGACAATGGAGCTCAGCTTGAAGGGCCT
GluThrAspThrGlyGlnValValTyrTyrAspAspAsnGlyAlaGlnLeuGluGlyPro

      1750       1760       1770       1780       1790       1800
GGCCAGGTTGGGTTGTCTTCCGACGGACCGAGCTACCAGATCAGCGTATACGAACGGGGA
GlyGlnValGlyLeuSerSerAspGlyProSerTyrGlnIleSerValTyrGluArgGly

      1810       1820       1830       1840       1850       1860
TTTCAGACGCCAGATTGGCTGAAGCACGCCGTGATCTACGAAATCATGCCGGATCGGTTC
PheGlnThrProAspTrpLeuLysHisAlaValIleTyrGluIleMetProAspArgPhe

      1870       1880       1890       1900       1910       1920
TACAATGGCAATATCGCCACGGAGGAGAATCCGAATACGCAAAAGGGGATTTATGTAGGG
TyrAsnGlyAsnIleAlaThrGluGluAsnProAsnThrGlnLysGlyIleTyrValGly

      1930       1940       1950       1960       1970       1980
GCCGATGGAACGGAGTCATTAGGCCCCATCCAGTTCCACGAGAACTGGGACTCGCCGCCC
AlaAspGlyThrGluSerLeuGlyProIleGlnPheHisGluAsnTrpAspSerProPro

      1990       2000       2010       2020       2030       2040
TATGATCCGAATATTCCTCCGTTATCTGATCCCAAAATTGCCAGTCTGCGAGGCAATGGC
TyrAspProAsnIleProProLeuSerAspProLysIleAlaSerLeuArgGlyAsnGly

      2050       2060       2070       2080       2090       2100
CAATGGAACATTGACTTTTTCGGAGGTGATTTGCAGGGCATCGAGGATAAGCTGGACTAC
GlnTrpAsnIleAspPhePheGlyGlyAspLeuGlnGlyIleGluAspLysLeuAspTyr

      2110       2120       2130       2140       2150       2160
CTGAAGAGCCTTGGAGTCAATACGCTGTATCTGATGCCCGTCTTTGAGGCGGAATCCAAT
LeuLysSerLeuGlyValAsnThrLeuTyrLeuMetProValPheGluAlaGluSerAsn

      2170       2180       2190       2200       2210       2220
CACAAATATGACACAGCCGACTATTTCAAGATTGACCCTGGATTTGGAACGCAGCAGGAC
HisLysTyrAspThrAlaAspTyrPheLysIleAspProGlyPheGlyThrGlnGlnAsp

      2230       2240       2250       2260       2270       2280
TGGCTGAATCTCGTACAGGCTGCGCACGCGAAGGGGTTCCATATCATTCTCGACGGGGTG
TrpLeuAsnLeuValGlnAlaAlaHisAlaLysGlyPheHisIleIleLeuAspGlyVal

      2290       2300       2310       2320       2330       2340
TTCGAAGATACCGGAAGTGACAGCGTATATTTCAACAAGTTCGGGAACTTCCACTCCAAC
PheGluAspThrGlySerAspSerValTyrPheAsnLysPheGlyAsnPheHisSerAsn
```

Fig. 3 - page 4 of 6

```
      2350       2360       2370       2380       2390       2400
GGTGCGTGGCAGGCGTACCTGAAGAACCAGCCGTCGCTGTCGCCCTACTACTCGTGGTAC
GlyAlaTrpGlnAlaTyrLeuLysAsnGlnProSerLeuSerProTyrTyrSerTrpTyr

      2410       2420       2430       2440       2450       2460
GTGTGGACAGGGAACACCTCAAACCCATACGATTCGTGGTTTCAGATCGACACGCTGCCA
ValTrpThrGlyAsnThrSerAsnProTyrAspSerTrpPheGlnIleAspThrLeuPro

      2470       2480       2490       2500       2510       2520
CTTACGGACACGTCGAACCCCGCCTATCAGCGATTCGTGTATGGGAGCGACAACTCAGTC
LeuThrAspThrSerAsnProAlaTyrGlnArgPheValTyrGlySerAspAsnSerVal

      2530       2540       2550       2560       2570       2580
GCGCGTGTGTGGATCCGGGAAGGTGCGGACGGATGGCGCTTGGACTCGGCCGACAACGGG
AlaArgValTrpIleArgGluGlyAlaAspGlyTrpArgLeuAspSerAlaAspAsnGly

      2590       2600       2610       2620       2630       2640
AATTTCAACACGGCATGGTGGGGTGGCTTTCGGCAGGCCGTGAAATCGATCGATCCCAAC
AsnPheAsnThrAlaTrpTrpGlyGlyPheArgGlnAlaValLysSerIleAspProAsn

      2650       2660       2670       2680       2690       2700
GCAGCGATCATCGGCGAGATCTGGGACAATGCGACGAATGACAATGGAACGGATTGGTTG
AlaAlaIleIleGlyGluIleTrpAspAsnAlaThrAsnAspAsnGlyThrAspTrpLeu

      2710       2720       2730       2740       2750       2760
ACGGGATCGACCTTCGACAGTGTAATGAACTACCAGTTCCGGAACGCCGTGATCGACTTC
ThrGlySerThrPheAspSerValMetAsnTyrGlnPheArgAsnAlaValIleAspPhe

      2770       2780       2790       2800       2810       2820
TTCCGCGGCACGTACAACGACGGAAACGTGCAGCACCACGCCGTCGACGCTGCGGGATTC
PheArgGlyThrTyrAsnAspGlyAsnValGlnHisHisAlaValAspAlaAlaGlyPhe

      2830       2840       2850       2860       2870       2880
AACCAGGAACTGATGCGCCTGTACAGTGAATATCCTCTGCAGTCGTTCTACTCGATGATG
AsnGlnGluLeuMetArgLeuTyrSerGluTyrProLeuGlnSerPheTyrSerMetMet

      2890       2900       2910       2920       2930       2940
AACCTTGTCGATTCGCAAGACACCATGCGGATCCTGACCATCTTGGAGAACGCGCCGCAG
AsnLeuValAspSerGlnAspThrMetArgIleLeuThrIleLeuGluAsnAlaProGln

      2950       2960       2970       2980       2990       3000
CCAGGCGATCTATCCGCGCTCCAGCAGGATGAGTACAGGCCGTCTCCTGCGGCTGAACAG
ProGlyAspLeuSerAlaLeuGlnGlnAspGluTyrArgProSerProAlaAlaGluGln

      3010       3020       3030       3040       3050       3060
TTGGGGATCGAGAGGCTGAAGCTTGTATCGGACTTTCAATTCAGCTTCCCGGGCGATCCG
LeuGlyIleGluArgLeuLysLeuValSerAspPheGlnPheSerPheProGlyAspPro

      3070       3080       3090       3100       3110       3120
ACCATCTTCTACGGCGACGAGGCAGGGCTCACTGGTTATTCGGATCCCCTCAATCGTCGG
ThrIlePheTyrGlyAspGluAlaGlyLeuThrGlyTyrSerAspProLeuAsnArgArg
```

Fig. 3 - page 5 of 6

```
      3130       3140       3150       3160       3170       3180
ACCTATCCGTGGGACAACCAGAATCTCGATCTCCTGAACCACTACCGCAAGCTCGGGGCC
ThrTyrProTrpAspAsnGlnAsnLeuAspLeuLeuAsnHisTyrArgLysLeuGlyAla

      3190       3200       3210       3220       3230       3240
ATTCGAAACGCCAATCCTGTGCTTCAGACGGGGGATTTCACGCCGTTGTACGCACAGGGC
IleArgAsnAlaAsnProValLeuGlnThrGlyAspPheThrProLeuTyrAlaGlnGly

      3250       3260       3270       3280       3290       3300
ATGGTGTACGCATTTGCAAGGACCATTCGGAATGGGCGAGATGTCTTCGGTGTGCCAGCG
MetValTyrAlaPheAlaArgThrIleArgAsnGlyArgAspValPheGlyValProAla

      3310       3320       3330       3340       3350       3360
GAGGATGCCACGGCCATTGTGGCGATCAACAATCAGAACCAAGCTATCACCGTGACCATT
GluAspAlaThrAlaIleValAlaIleAsnAsnGlnAsnGlnAlaIleThrValThrIle

      3370       3380       3390       3400       3410       3420
CCGACGGATGGGACGGTTGCGGACGGGTCCACGATGCTCGATGAACTGAACAACCAGTGG
ProThrAspGlyThrValAlaAspGlySerThrMetLeuAspGluLeuAsnAsnGlnTrp

      3430       3440       3450       3460       3470       3480
TACAAGGTGCAGAATGGTGGCATTACACTCACGCTGCAATCGTATCAAGGTGCCATTTTG
TyrLysValGlnAsnGlyGlyIleThrLeuThrLeuGlnSerTyrGlnGlyAlaIleLeu

      3490       3500       3510       3520       3530       3540
GTGACGCCGAGCGACGCGCCGATGGCTTATCTGCAAGAGGAGGATTCTCAGAACGAGATT
ValThrProSerAspAlaProMetAlaTyrLeuGlnGluGluAspSerGlnAsnGluIle

      3550       3560       3570       3580       3590       3600
GCGTGGACGCCTGTGCAAGGTGCCATCGGTTATCGCGTCTGGAGACAGAATCCGAATGGA
AlaTrpThrProValGlnGlyAlaIleGlyTyrArgValTrpArgGlnAsnProAsnGly

      3610       3620       3630       3640       3650       3660
CAATGGGTGCCCTTTGGACCTGTGCTTCCTGCCACGGACTTGAGTGTCACGGTGGAACGC
GlnTrpValProPheGlyProValLeuProAlaThrAspLeuSerValThrValGluArg

      3670       3680       3690       3700       3710       3720
GATGCATATGCGCAAACGTTTGCTGTACAAGCGCTGTTTTCGGCGTCTGATCACGCCCAG
AspAlaTyrAlaGlnThrPheAlaValGlnAlaLeuPheSerAlaSerAspHisAlaGln

      3730       3740       3750       3760       3770       3780
TCTCCGGTGTCGGCACCTAAGACGGTATCGCTTCCCGTCGATGTGCCCGCGGTACGCCTG
SerProValSerAlaProLysThrValSerLeuProValAspValProAlaValArgLeu

      3790       3800       3810       3820       3830       3840
AGTCAGCCGATCGTTAGTGGTCGTGTGGTTGGAGATCGTGCGATGGTCTCGATCACGCCG
SerGlnProIleValSerGlyArgValValGlyAspArgAlaMetValSerIleThrPro

      3850       3860       3870       3880       3890       3900
GTTTCAGGCGCGACGCAGTATGTGATCTACCAGAGACAGGGCGACGGATCGTATGCTCCG
ValSerGlyAlaThrGlnTyrValIleTyrGlnArgGlnGlyAspGlySerTyrAlaPro
```

Fig. 3 - page 6 of 6

```
      3910       3920       3930       3940       3950       3960
GTCGCGACGGTCTCCACAAGTGGCGATTCCGCAGCTATAGGGGAAGTTCCTGCGCAAGGT
ValAlaThrValSerThrSerGlyAspSerAlaAlaIleGlyGluValProAlaGlnGly

      3970       3980       3990       4000       4010       4020
CCGGCCAACTCGCCTCACGCGACGATTCGCGTGACAGTGCCCGTACCTGCAGGTTTCTCG
ProAlaAsnSerProHisAlaThrIleArgValThrValProValProAlaGlyPheSer

      4030       4040       4050       4060       4070       4080
TCGGTGACCTACCGCGTGGCTGCGCAAAACGAAGATGGGCAAGCTGTGACCAATCCATTG
SerValThrTyrArgValAlaAlaGlnAsnGluAspGlyGlnAlaValThrAsnProLeu

      4090       4100       4110       4120       4130       4140
ACCCTATCGCTCTCGAAAAAGTGATGCCTCGCGAAAAGGGCTATCGGAATTTTTTCGAGA
ThrLeuSerLeuSerLysLys---
```

# INTERNATIONAL SEARCH REPORT

International Application No PCT/FI 92/00138

## I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/56; C12N1/21; C12N9/28; A23K3/03

## II. FIELDS SEARCHED

Minimum Documentation Searched[7]

| Classification System | Classification Symbols |
|---|---|
| Int.Cl. 5 | C12N |

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

## III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

| Category[10] | Citation of Document,[11] with indication, where appropriate, of the relevant passages[12] | Relevant to Claim No.[13] |
|---|---|---|
| X | EP,A,0 410 498 (GIST-BROCADES) 30 January 1991<br>see the whole document | 1 |
| X | AGRICULTURAL AND BIOLOGICAL CHEMISTRY.<br>vol. 50, no. 1, 1986, TOKYO JP<br>pages 23 - 31;<br>M. KANNO: 'A Bacillus acidocaldarius alpha-amylase that is highly stable to heat under acidic conditions'<br>cited in the application<br>see the whole document | 1 |
| X | STARCH<br>vol. 31, no. 5, 1979, WEINHEIM DE<br>pages 166 - 171;<br>E. BOYER ET AL: 'Isolation and characterization of unusual bacterial amylases'<br>cited in the application<br>see the whole document | 1 |

-/--

[10] Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

## IV. CERTIFICATION

| Date of the Actual Completion of the International Search | Date of Mailing of this International Search Report |
|---|---|
| 29 JULY 1992 | 05. 08. 92 |
| **International Searching Authority** | **Signature of Authorized Officer** |
| EUROPEAN PATENT OFFICE | VAN DEN SCHAAL C.A. |

3

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

| Category* | Citation of Document, with indication, where appropriate, of the relevant passages | Relevant to Claim No. |
|---|---|---|
| A | WO,A,8 901 970 (PLANT GENETIC SYSTEMS) 9 March 1989<br>see examples 1,2,7-25 | 1-9,11, 12,14-18 |
| A | GENE.<br>vol. 19, 1982, AMSTERDAM NL<br>pages 81 - 87;<br>I. PALVA: 'Molecular cloning of alpha-amylase gene from Bacillus amyloliquefaciens and its expression in B. subtilis'<br>see the whole document | 1-8,10, 11,13-15 |

# ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. FI 9200138 SA 59061

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 29/07/92

| Patent document cited in search report | Publication date | Patent family member(s) | | Publication date |
|---|---|---|---|---|
| EP-A-0410498 | 30-01-91 | AU-A- | 5953890 | 17-01-91 |
| | | WO-A- | 9100353 | 10-01-91 |
| | | JP-T- | 4500756 | 13-02-92 |
| WO-A-8901970 | 09-03-89 | EP-A- | 0311469 | 12-04-89 |
| | | JP-T- | 2501266 | 10-05-90 |